VOL. 13 No. 9, pp. 237-268 MARCH 1952

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

DELTA MODULATION, A NEW MODULATION SYSTEM FOR TELECOMMUNICATION

by J. F. SCHOUTEN *), F. de JAGER and J. A. GREEFKES.

621.396.619.16

For the transmission of information such as speech or music the aim is to design a system possessing the optimum properties, having regard to the properties of the transmitting channel on the one hand and, on the other hand, the characteristics of the signal to be transmitted. Recently systems have been developed in which a quantization of the signal is applied both in amplitude and in time. Such a system is described in this article. Its essential feature is the employment of a special kind of negative feedback. Compared with other quantizing modulation systems the apparatus has been greatly simplified.

Signal and interference

A remarkable development is taking place in the technique of telecommunications, in that speech and music are being transmitted by means of a pattern of unit signals, such as already known in telegraphy.

The desire for such a drastic departure from the usual methods of transmission, such as amplitude modulation for instance, has arisen from the fact that it is becoming more and more difficult to ensure interference-free reception at very great distances. In modern carrier-telephony systems the signal transmitted by cable over a distance of 100 km is attenuated 240 db, i.e. a power ratio of 10^{-24} times. This attenuation, which is in the order of the ratio of the power emitted by the sun to that of a pocket-torch lamp, does not in itself constitute any difficulty: in transcontinental communications such attenuations may reach some tens of thousands of decibels, but they are compensated by means of a series of repeaters (relay stations) installed at intervals along the route. But, besides being attenuated, the signal is also distorted. The distortion may consist, for instance, in noise, crosstalk from neighbouring channels, and variations in the level of the signal. These effects, in themselves only small, act along the whole transmission circuit, so that the ultimate distortion of the signal received is a cumulation of very small effects.

When a table is covered with dust it can be cleaned by brushing the dust off, for we know what is the table and what is the dust. However, at the receiving end of the communication system referred to it is impossible to distinguish between signal and interference. This means that the "transmission dust" cannot be brushed off, and so with the usual systems of communication cumulation of interference is unavoidable.

There are, however, communication systems where the interference can indeed be distinguished from the signal, and with such systems, when intermediate repeater stations are used, a practically interference-free signal can be received even at very great distances. Before dealing with these systems we shall discuss some other modern methods of transmission which have, as it were, pointed the way to achieving the ultimate object of suppressing interference. These methods are illustrated in *fig. 1.*

How cumulation of interference is counteracted

Amplitude modulation on a carrier (fig. 1*b*) of the signal to be transmitted (fig. 1*a*) can be replaced by amplitude modulation on a series of equidistant pulses: pulse-amplitude modulation (fig. 1*c*). An audio or video signal is given as

*) Of Philips' Telecommunication Industries, Hilversum.

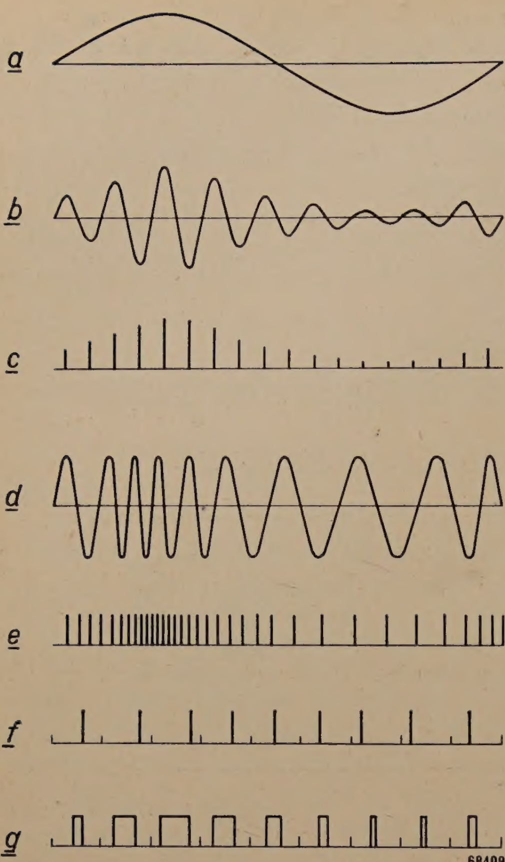


Fig. 1. Diagrammatic representation of various non-quantized modulation systems. *a*) The modulating signal. *b*) Carrier amplitude modulation. *c*) Pulse-amplitude modulation. *d*) Carrier frequency modulation. *e*) Pulse-frequency modulation. *f*) Pulse-position modulation (note the marked time intervals within which each pulse can move). *g*) Pulse-width modulation.

a continuous function; the time is the independent variable, while the dependent variable is the sound pressure, respectively a voltage at the pick-up. It appears that if the value of this voltage is known at equidistant instants the total signal, i.e. including all instantaneous values between those instants, can be wholly reconstructed, provided the frequency bandwidth of the signal is not greater than half the rate of sampling. Thus, the wider the range covered by the frequencies of the signal to be transmitted, the greater is the required sampling rate.

In pulse-amplitude modulation this principle is applied directly. In fact this method is analogous to carrier-amplitude modulation; only the continuous carrier is replaced by discrete pulses. For a frequency band of, say, 3400 c/s to be transmitted, a pulse-repetition frequency of about 8000 c/s is employed. As far as the separation of signal and interference is concerned this method does not, however, offer any advantage over carrier modulation.

Almost simultaneously with pulse-amplitude modulation, frequency modulation (fig. 1*d*) made its appearance in telecommunication technique. With this method it is not the amplitude but the frequency of the carrier that is modulated. Here, too, the continuous carrier can be replaced by a series of pulses, so that pulse-frequency modulation (fig. 1*e*) is obtained [1]). A special form of this is pulse-position modulation (fig. 1*f*), also called pulse-time modulation, whereby the signal is built up from pulses all having the same amplitude and width, called unit pulses. The information is contained in the time pattern of these pulses, such that each pulse occupies a certain place in the time interval allotted to it, according to the instantaneous value of the signal [2]). This offers an entirely new possibility of counteracting interference: if the effects of interference in the transmission channel are sufficiently limited so that they do not severely mutilate the unit pulses then the pattern is retained and can be restored by regenerating the partly mutilated pulses at the receiving end.

Still it is not possible to reach a complete separation between signal and interference, so that some interference remains and is cumulative. The interference in the transmission channel acts in the first instance on the amplitude of the signal transmitted. If the pulses were quite rectangular in shape complete separation of the signal from interference would be possible, provided the interference does not exceed a certain value. Since, however, the slope of the edges of the pulses always has a finite value, there is still some variation in the time at which a pulse is received, i.e. in the position of the pulse in its interval. The resultant interference cannot be separated from the signal and so there is again a cumulative effect.

There is much less interference with pulse-position modulation than with pulse-amplitude modulation, and there is much less interference in the case of carrier-frequency modulation than in that of carrier-amplitude modulation. Frequency modulation and pulse-position modulation have led to the fundamental conception that it is possible to reduce the effect of interference by increasing the frequency band of the high-frequency signal transmitted. This possibility, however, is not efficiently utilised

[1]) Pulse-frequency modulation is something that has existed in nature for millions of years, viz. in the transmission of signals via the nerves of animals. When a nerve is excited a series of electrical pulses of equal amplitude are propagated along the nerve; the number of pulses per second (up to some hundreds) is a measure for the strength of the stimulus.

[2]) Pulse-position modulation has been mentioned in this journal before; see C. J. H. A. Staal, An installation for multiplex-pulse modulation, Philips Techn. Rev. 11, 133-144, 1949.

in pulse-position modulation, nor in its modified form of pulse-width modulation (based on the principle that the width of the pulses is varied according to the instantaneous value of the signal, fig. 1g).

Pulse-position modulation with the aid of rectangular pulses would, as already stated, give complete suppression of interference, but the frequency band required for undistorted transmission of these signals is infinitely large.

Quantization of the signal

For pulse-position modulation unit signals are employed. This implies a sort of quantization: only discrete values (in this case only the unit value) of the pulse amplitude are permitted. The information is determined, however, by a continuously variable pattern of these pulses, in the sense that any particular pulse may be at any point within the continuous time range allotted to it. One can now go a step further and transmit only discrete values — say, the multiples of a certain unit — of the continuous scale of instantaneous values of the signal. Something similar is regularly done in everday life when expressing length in kilometres, metres or millimetres, temperature in degrees or tenths of a degree, time in hours, minutes or seconds. The consideration behind this inaccurate quotation is the fact that in practice a limited accuracy, depending on the object in view, always appears to be sufficient.

The steps in which the instantaneous value of an audio signal is measured can likewise be chosen with a sufficient fineness to obtain a certain quality of reproduction. The loss in fidelity is offset by the advantage gained in being able to eliminate completely the active effects of interference, provided it is kept below a certain limit.

This can be illustrated with the following example, for which we shall revert to the method of pulse-position modulation. Complete quantization of this modulation means that in the time interval allotted to it each pulse can occur only at a discrete number of fixed places, as illustrated in *fig. 2.* If the interference brings about a displacement of the pulse positions which is less than half the difference in time between two of the discrete places then, in spite of this interference, at the receiving end one knows exactly what pulse position, thus what signal, is intended. The interference can therefore be entirely separated from the signal and completely eliminated, provided it does not exceed a certain threshold value.

Direct application of the quantization process to this or to one of the other modulating systems mentioned thus gives the desired suppression of interference. But, as will be shown later, from the point of view of efficiency of the transmission system it is uneconomical to proceed in this way.

Special methods of modulation have therefore been developed for transmission of quantized signals. One of these methods is that of delta modulation, which will now be discussed.

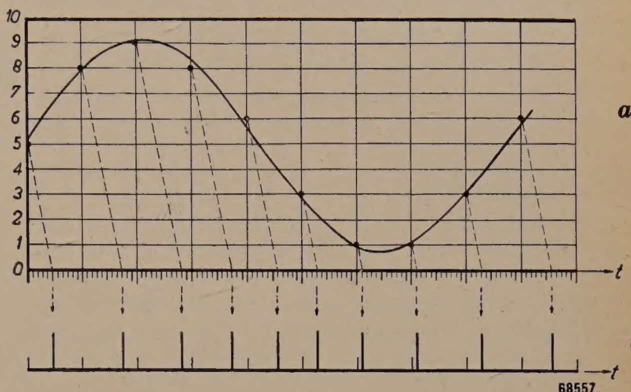


Fig. 2. Quantization of pulse-position modulation. *a*) The instantaneous signal value at the beginning of each interval is approximated by a quantized value. This is "translated" into the series of transmitted pulses (*b*) by a corresponding pulse position in the interval. In this diagram the quantization is carried out in 11 steps. Thus for each pulse there are 11 discrete positions available in its interval. This system is not applied in practice because it is not effective.

Delta modulation

Principle

Delta modulation is a pulse modulation system, pulses being produced at the sending end at equidistant instants. Not all pulses are, however, transmitted. The question whether a pulse is to be transmitted or not is decided in the following way. An "echelon curve" is formed from the pulses in a manner to be explained later. This echelon curve is compared with the modulating signal (*fig. 3b*). If at the moment when a new step has to be added the value of the modulating signal is greater than that of the step-shaped signal, then a positive step is added, but if the value of the modulating signal is smaller a negative step is formed. Thus the echelon curve oscillates and approximates to the modulating signal, being the quantized form thereof. The formation of a positive step is arranged to initiate the transmission of a pulse, whereas the formation of a negative step is not and results in a gap in transmission at that moment in time. What is then transmitted is the series of pulses shown in fig. 3*a*.

At the receiving end the step-shaped approximating signal is built up again from the series of pulses received and finally converted into a con-

tinuous signal giving the reproduction of the original modulating signal. It is clear that if the interference in the transmission channel does not mutilate or displace the pulses as such beyond recognition then this interference can be completely eliminated.

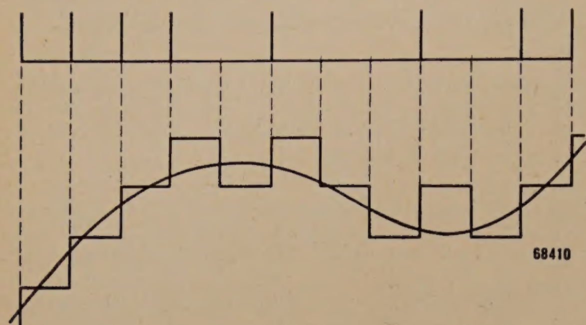


Fig. 3. Principle of delta modulation. The series of pulses emitted by the transmitter (a) is generated after comparison of the modulating signal (the continuous curve in b) with a step-like approximating signal with equidistant steps built up in the transmitter. When the approximating signal is smaller than the modulating signal a pulse is emitted, but if it is larger the pulse is suppressed. At the receiving end the approximating signal is formed from the series of pulses by integration and from that signal the modulating signal is reconstructed by smoothing out the steps.

This is only a very sketchy exposition and in practice the method is not so simple as that, but anyhow it gives the essence of the method, which is recapitulated in the following concise form.

With the system of delta modulation, from a series of equidistant pulses transmitted, which may have the value 1 or 0, at the receiving end a signal is built up which has to form the nearest possible approximation of the original modulating signal. This approximating signal is constructed not only at the receiving end but also at the transmitter. At the sending end the original signal can therefore always be compared with an approximation similar to that which would be obtained with the aid of the quantized pulses at the receiving end. If the approximating signal is smaller than the input signal then a pulse is sent out by the transmitter, but if the approximating signal is larger than the input signal the next pulse is suppressed.

The "auxiliary receiver" at the sending end, which supplies the approximating signal, is in its simplest form an integrating network. The output voltage from this network is compared with the original signal in a difference meter, a simple resistance network, followed by an amplifier. Thus a sort of inverse feedback is applied in the transmitting circuit. This feedback, however, differs from the usual inverse feedback in two respects. First, the feedback circuit is closed only at the moments when it has to be decided whether a pulse is to be transmitted or not. Second, the voltage acting on the feedback circuit is not proportional to the difference voltage but dependent only upon the polarity of the difference voltage. This inverse feedback is virtually quantized in position and amplitude.

In *fig. 4* a block diagram is given representing the circuit of the transmitter. We see here a pulse generator, *I.G.*, which yields a series of equal pulses with a repetition frequency of, say, 60 000 or 100 000 pulses per second. These pulses have to pass through a pulse modulator, *I.M.* If the voltage at the point *2* is negative then *I.M.* gives a pulse having the same sign as the incoming pulse; if, on other hand, the voltage at *2* is positive then a pulse with the opposite sign is sent to point *3*. The pulses coming from the pulse modulator *I.M.* are sent to an integrating network *I* (i.e. the "auxiliary receiver"). Here each positive pulse will raise the output voltage a certain amount and each negative pulse will lower it by the same amount. In this way an echelon curve or approximating signal *4* is obtained. In the difference meter *V.M.* this signal is compared with the modulating signal *5*, the output voltage always indicating the difference between the two signals and thus controlling the pulse modulator. The next pulse is therefore positive when the approximating voltage at that instant is lower than the input signal and negative when it is higher. The series of positive and negative pulses generated in this manner are

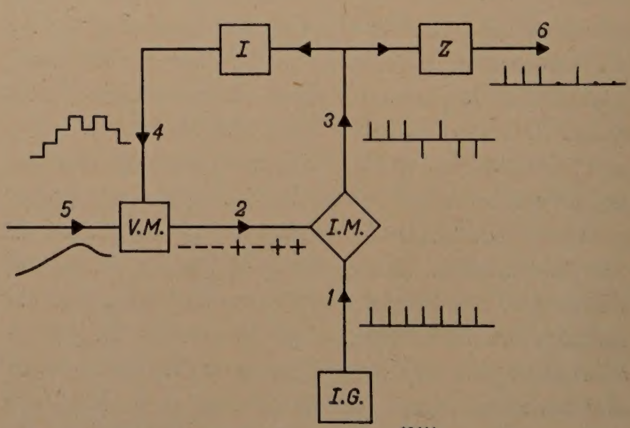


Fig. 4. Block diagram of the transmitter. The pulse generator *I.G.* supplies a series of identical equidistant pulses *1* which are passed on to the transmitter by the pulse modulator *I.M.* either unchanged or with opposite polarity. The action of the pulse modulator is governed by the polarity of the difference voltage *2* obtained in the difference meter *V.M.* by comparison of the modulating signal *5* and the step-like approximating signal *4* formed from the series of pulses *3* by integration in *I* ("auxiliary receiver"). The transmitter suppresses the negative pulses and passes the signal *6* to the aerial or the transmission line.

passed not only to the auxiliary receiver but also to a transmitter Z (for instance a radio relay system) which suppresses the negative pulses and transmits the positive ones.

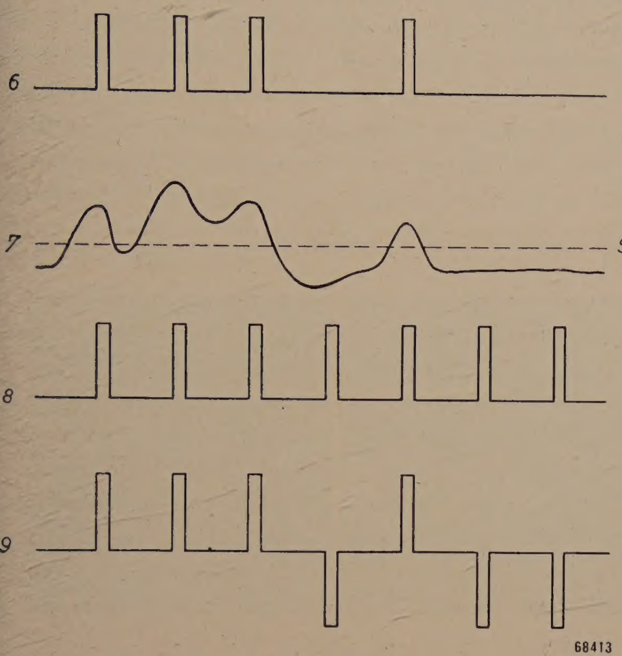


Fig. 5. The series of pulses *6* emitted by the transmitter of fig. 4 becomes distorted in the course of transmission, assuming, for instance, the shape of the signal *7*. This signal influences a pulse modulator in the receiver in such a way that, provided the pulse amplitude exceeds a certain limit *s* (mostly chosen equal to half the original pulse amplitude), the pulses *8* produced by a pulse generator are passed on with the correct polarity. If the signal *7* lies below the level *s* the pulses produced by the generator are replaced in the pulse modulator by pulses of the opposite polarity. In this way the series of pulses *9* is obtained.

Fig. 5 shows to what changes the pulses on their way through the aether are subject. The pulse pattern as transmitted corresponds to *6*, while owing to the interference the pattern as received is more or less distorted, resembling, for example, that depicted by *7*. Now in order to reconstruct from this pattern the step-shaped approximating signal, in the receiver (*fig. 6*) a pulse modulator is employed which in principle functions in the same way as the pulse modulator in the transmitter: the pulse (*8*) coming from a pulse generator *I.G.* is passed through unmodulated if the incoming pulse exceeds a certain value *s* (for which, as a rule, half the height of the undistorted pulse is taken), but if the incoming pulse is smaller than that value then the pulse modulator *I.M.* gives a pulse of the opposite sign. At the output of the pulse modulator a combination of pulses (*9*) is obtained which is identical to the signal *3* (see figs 4, 5 and 6). (The pulse generator has to be synchronized with the incoming pulses, the apparatus required for this having been omitted from the diagram in fig. 6 for the sake of simplicity.)

When this series of pulses is applied to an integrating network, *I*, the same step-shaped signal (*10*) is obtained as was generated at the transmitter (signal *4*). By passing this signal through a low-pass filter *F* the steps are smoothed out into a continuous signal *11* closely resembling the original input signal.

In this manner interfering signals of at most half the pulse amplitude can be permitted in the course of transmission without being perceptible in the signal.

Sound reproduction by the delta modulation system

Obviously the step-shaped approximating signal (and thus also the continuous output signal derived from it) will always show some small deviations with respect to the original signal. If the modulating signal is sinusoidal then at the receiver a sinusoidal signal is obtained against a background of noise, the so-called quantizing noise, due to the approximation of the instantaneous values of the speech signal by discrete values. The higher the rate of sampling, however, the less is this quantizing noise. In practice, with a pulse frequency of 40 000 pulses per second there is still a perceptible grating quality in the speech signal transmitted, but with a pulse frequency of 100 000 pulses per second there is practically no noticeable quantizing noise. Due to this high pulse frequency a much wider band is required for the transmission than is needed in the case of direct modulation of the signal on a carrier, since the bandwidth required is at least

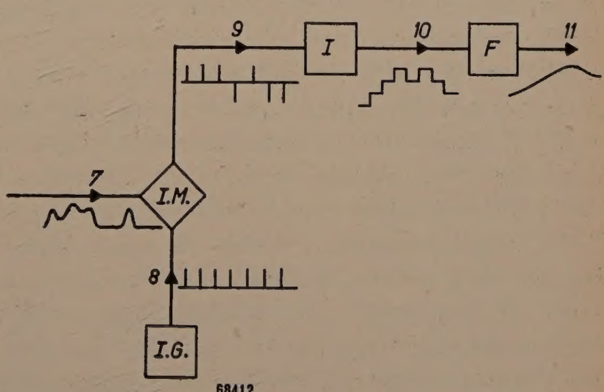
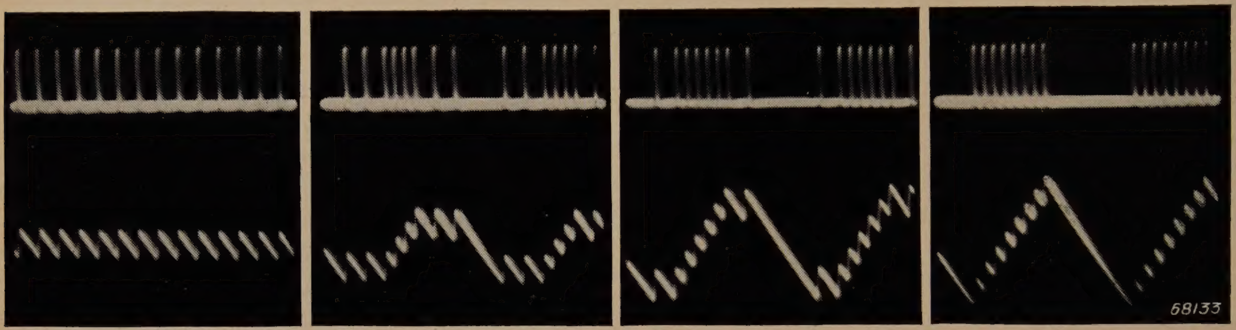


Fig. 6. Block diagram of the receiver. The distorted signal *7* influences the pulse modulator *I.M.* in the manner described under fig. 5. The pulses *8* produced by the pulse generator *I.G.* may be passed on with the correct or opposite polarity and the signal *9* thus produced is converted by an integrator *I* into the echelon curve *10*. From this the continuous-wave signal *11* is formed in the low-pass filter *F*, this signal being the reproduction of the original signal at the input of the transmitter.

equal to half the repetition frequency of the pulses. From the physical point of view this is to be regarded as the price to be paid for the suppression of the interference. Especially for radio relay systems, however, this does not constitute any serious objection, since in this frequency range it is easier to increase the bandwidth than to improve the signal-to-noise ratio.

In practice the pulse modulator is so arranged that only pulses of the values 1 and 0 appear at the output. The input signal is then more or less approximated by a sawtooth instead of an echelon. maximum amplitude has been reached that can be observed at the receiver. Larger input amplitudes will not further modify the approximating signal, and thus at the receiver the same output signal will be obtained. It therefore appears that with delta modulation there is a limitation of the amplitudes to be transmitted. This limitation is reached when the signal voltage assumes a certain slope corresponding to the permanent emission of positive pulses. The higher the frequency of the input signal, the smaller is the maximum amplitude of the output signal, as will be shown below.



a b c d

Fig. 7. The sawtooth-shaped approximating voltage and the corresponding series of pulses transmitted, for an A.F. sinusoidal input signal of 4000 c/s. The rate of sampling of the pulses supplied by the generator is 64 kc/s. *a*) The amplitude of the sine curve is zero; the pulses emitted are alternately 1 and 0. *b*) and *c*) The amplitude is gradually increased, as a result of which the pattern of the series of pulses is changed. *d*) The amplitude has by now become so large that the system reaches complete modulation, and even overmodulation. Even larger amplitudes will not produce any other pulse pattern.

This does not, however, alter the principle of the system.

In *fig. 7* some oscillograms of such an approximating signal are reproduced, together with their corresponding series of pulses. These approximating signals built up from unit pulses are smoothed out at the receiver with the aid of a low-pass filter, as already mentioned, and then give a reproduction of the (in this case sinusoidal) input signal. The frequency of the input signal in the case under consideration is 4000 c/s, the rate of sampling is 64 000 pulses per second. From the four oscillograms it is to be seen how the configuration of the pulses changes with increasing amplitude of the signal transmitted. In fig. 7*a* the input signal is zero. The pulses transmitted are then alternately 1 and 0, and since the sawtooth-shaped approximating signal then has the fundamental frequency of 32 000 c/s, thus containing no frequencies lying within the pass band of the filter, the value of the output signal behind the filter in the receiver will likewise be zero. In figs 7*b* and 7*c* the amplitude has been gradually increased, while in fig. 7*d* the

Other methods for transmitting quantized signals

From the foregoing it is seen that with the method of delta modulation two principles are applied: integration of pulses, and application of a special type of inverse feedback. Owing to the inverse feedback the essence of the method lies in the transmitter responding only to what has been changed with respect to the earlier situation. Each pulse transmitted is only a correction of the preceding signal.

Another system of modulation based on the use of quantized signals is that of pulse code modulation [3]). This system works as follows. As in the case with delta modulation, only the presence or absence of a pulse is of importance. To be able to distinguish more than two amplitude levels with the aid of this "all or nothing" principle, the instantaneous values of the message signal are denoted not by one pulse (present or not) but by a code group of 5 or 7 equidistant pulses each of

[3]) See, e.g., W. M. Goodall, Telephony by pulse code modulation, Bell System Technical Journal **26**, 395-409, 1947.

which may have the value 1 or 0. The position of a pulse in the code group determines its significance. This position principle is met with also in ordinary numerical systems (it was first employed systematically by Indian mathematicians about the year 600). In the decimal system all numbers can be denoted with the aid of only 10 symbols, the position of a symbol in a number determining its value (units, tens, hundreds, etc.). With the method of pulse code modulation, instead of the decimal scale, the simpler binary system is used. The position of a pulse in the code group determines whether it will have the value 1, 2, 4, 8, and so on. In this way, with code groups of five pulses, 32 different signal values can be denoted. This is illustrated in *fig. 8*.

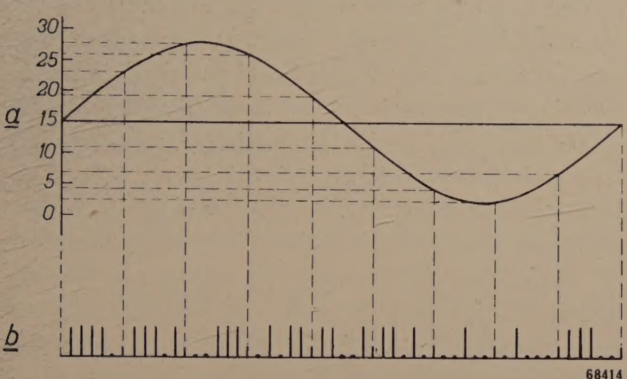


Fig. 8. Principle of pulse code modulation with code groups of 5 pulses. At the beginning of each time interval corresponding to a code group the value of the signal to be transmitted (a) is rounded off to one of the discrete values 0 to 31 and "translated" into a pulse pattern (b). This pattern is formed by the presence or absence of pulses at the five places marked by dots within the group. A binary system is applied, each place corresponding to a binary value and the measured amplitude thus being expressed according to the binary scale. From left to right the five points in a code group then represent 1, 2, 4, 8 and 16 amplitude "units".

Delta modulation may be regarded as a modified form of pulse code modulation in so far as both systems employ a coding of the signals to be transmitted. Pulse code modulation employs 5 or 7 coding units (sometimes even more), whilst delta modulation, at least in the simple form described here, employs a coding with the aid of one unit, viz. one step in the echelon curve. This is possible on account of the fact that with delta modulation only corrections of the signal are transmitted, whereas in the case of pulse code modulation the complete signal values are transmitted.

With the introduction of pulse code modulation a binary five-unit code came to be applied in telecommunications for the fourth time. In 1837 Cooke and Wheatstone built in London (between Euston and Camden Town) a transmission system of five copper wires connected at the receiving end to a telegraph instrument with five needles registering twenty different letters. In 1926 Rainey obtained a patent for a code system for picture telegraphy which may be regarded as the precursor of the present-day pulse-code modulation systems. The modern telex system, the principle of which was invented by Baudot in 1875, employs a code system usually of 5 units replacing the rather peculiar Morse code of telegraphy (the peculiarity about the Morse code is that the length of the signs is adapted to the frequency in which the signs occur, and therefore has a particular significance).

The systems can be compared only by means of the basic theory concerned with the transmission of information where the concept of information is given a well-defined meaning. This would lead us too far afield, so we shall mention only some of the results reached from the comparison and include in it also the (hypothetical) quantized pulse-position modulation as described on page 239. There are two kinds of interference to be distinguished, the interference which is inherent in the system itself and that which is introduced in the course of the transmission. In the case of non-quantized systems, in theory only a negligible interference (noise) arises in the transmitter and in the receiver, so that if the transmitter and the receiver were directly interconnected the reproduction would be quite free of interference. But, as already pointed out, along the path of transmission there is some interference which is cumulative. If this interference in the course of transmission remains below a certain threshold value then the transmission interference can be entirely eliminated by quantization. As we have seen, quantization means that an "interference" — the quantizing noise — is introduced in the transmitter itself, but this interference does not take part in the cumulation. Moreover, the quantizing noise can be made as small as desired by increasing the band covered by the high-frequency signal transmitted. The transmission interference, however, cannot be reduced.

Quantized pulse-position modulation is characterized by the fact that the threshold value of the transmission interference below which that interference is inactive is relatively low. If the pattern is not to be lost, the pulses may be displaced only slightly in the interval allotted to them. Further it can be shown mathematically that in this system the quantizing noise decreases inversely with increasing bandwidth f.

In coded systems, such as delta modulation and pulse code modulation, it is just the coding that allows of a much greater transmission interference. Furthermore in these two systems the quantizing noise appears to decrease far more with increasing bandwidth, according to a law of $f^{-5/2}$

and of approximately 2^{-f} respectively. This clearly shows the inefficiency, already mentioned, of quantized pulse-position modulation. The same inefficiency is found in all other quantized but not coded systems.

The reason why delta modulation and pulse code modulation are both more advantageous from the point of view of interference suppression than quantized, non-coded, normal systems lies in the use of inverse feedback in the system first mentioned and in the employment of more coding units in the second system. However, also in delta modulation a coding in more units can be applied, for instance by not only judging the difference measurement according to the sign but also dividing the difference into a number of steps and expressing it in the pulse picture transmitted. In this manner the signal-to-noise ratio can be improved for the same bandwidth. It can be shown that in order to obtain the same quality of reproduction fewer units are needed when inverse feedback is applied than when such is not employed.

Properties of delta modulation

Let us now return to the system of delta modulation described. From what has been said above it follows that from the point of view of quantizing noise delta modulation is less advantageous, for the same bandwidth, than pulse code modulation. It appears that in order to obtain a certain quality in the reproduction of speech in the case of delta modulation the band has to be about 50% wider than that required for the same quality with pulse code modulation. On the other hand, however, the apparatus is very much simpler. Not only is there no coding device required, but, since with delta modulation all the pulses are of equal significance, the requirements to be met for the circuit tolerances are less severe; in the difference meter only the polarity of the voltage is determined, so that the amplifiers for the difference voltage need not possess strictly linear characteristics. It is also permissible for a valve adjustment, or for the amplitude of the pulse, to change gradually with time, because in that case the inverse feedback ensures that the signal is not distorted. Finally, the construction of the feedback and integrating networks in the receiver is not very critical; deviations do not give rise to non-linear distortion (as is the case with the networks employed for pulse code modulation) but change only slightly the frequency characteristic of the whole system, which change can be corrected by a simple egalization. In the designing of the complicated quantization apparatus required for pulse code modulation great care has to be taken to ensure that it is linear, whereas for the quantization in the case of delta modulation a simple and essentially linear network is employed.

It would lead us too far to discuss the technical details of the circuit, but mention is to be made of the fact that the frequency characteristic of the network in the feedback circuit can be given an entirely different form for the frequencies above the audio range than for the lower frequencies, thereby reducing the quantizing noise.

Both with pulse code modulation and with delta modulation the possibilities are limited: the signals to be transmitted must not be too large, while at the receiver very small signals cannot be distinguished from the zero signal. In the case of pulse code modulation this limitation is governed by the size of the steps in the 5- or 7-unit code, the smallest signal to be transmitted corresponding to a code group with the "signal value" 1, and the largest corresponding to the transmission of a code group with the signal value 32 (5-unit code) or 128 (7-unit code). Thus the limitation is independent of frequency.

As already seen from fig. 7*d*, the limit in the case of delta modulation, for the large signal values, is reached when the signal voltage has a certain slope corresponding to the permanent transmission of positive pulses. If, however, the signal voltage is so small that notwithstanding the presence of the signal the approximating voltage retains the shape depicted in fig. 7*a*, and thus the pulses are alternately transmitted and suppressed, then nothing is to be perceived of the signal behind the low-pass filter in the receiver. By means of special circuits the smallest signal to be reproduced (threshold value) can be made somewhat smaller than what follows from this reasoning, but the minimum amplitude that can be observed at the receiver always remains finite.

As far as reproduction of the sound is concerned, the important consequence of the signal amplitude being limited owing to the existence of a maximum slope in the case of delta modulation is that the maximum amplitude that can be transmitted is inversely proportional to the frequency. Since in the case considered the threshold value of the signals received does not depend upon the frequency, this means that the dynamic range, i.e. the ratio of the maximum to the minimum amplitudes at the receiving end, becomes smaller as the frequency is raised.

This deviation from the ideal reproduction of the audio signal happens to coincide with a property inherent in the reproduction of speech and music and in the receptive qualities of the human ear, so that as far as its physiological effect is concerned it does not adversely influence the transmission: the maximum amplitude of the sound pressure in the audio spectrum decreases with increasing frequency, as is likewise the case with the dynamic range perception of the ear (especially at high frequencies).

A third quantity, one that does not occur in normal acoustics, is the gamma of the transmission system, by which is to be understood the number of distinguishable steps between the maximum and minimum amplitudes transmitted. With delta modulation the gamma becomes smaller as the frequency increases, and although physiological experiments have shown that in the range of speech frequencies this should lead to a deterioration in the quality of reproduction, in practice this is not evident.

For frequencies above 6000 c/s there is a reduction of the gamma with increasing frequency also in hearing. At these frequencies delta modulation therefore matches the properties of the human ear in two respects.

This brings us to the following highly important point. An optimum transmission system is characterized by the fact that it transmits what is to be transmitted and nothing more than that. When the signal power decreases as a function of the frequency then any possibilities offered by the system for the transmission of large powers at high frequencies are superfluous and undesired. When, as appears to be the case, the human ear perceives smaller dynamic ranges at the higher frequencies, in the optimum system it would be impossible for larger dynamic ranges than desirable to be transmitted. This explains why in the case of delta modulation a relatively small bandwidth suffices. The matching of this system to the peculiarities both of the speech signal and of the human ear partly neutralizes the mathematical deviations from the ideal reproduction.

The deeper significance of delta modulation

Being in essence based upon inverse feedback, delta modulation belongs to the family of regulating systems and servo-mechanisms, but with its quantization both in amplitude and in time it is a member that has been little studied.

As already stated, only corrections of the preceding signal are transmitted, this being brought about with the aid of the integrating network. Mathematical integration, however, would mean that even the distant past would contribute towards the present amplitude. Considering that there is a definite limit set to the time within which earlier signal values of a speech or music signal are related to the instantaneous value, those earlier values may gradually be forgotten. A resistance-shunted capacitor (as is present in our integrating *RC* network) thus supplies the first useful variation of the purely mathematical integration. A second variation consists in the fact that by adding more elements to the network an attempt is made to make the output signal correspond to the most probable continuation of the signal to be transmitted (that continuation where the first or more time derivatives of the signal voltage do not change with time). All the transmitter need do is then to transmit, via the difference meter, the corrections with respect to this most probable continuation.

This can also be formulated in another way. Each audio or video signal — whether speech, music or picture — is characterized by certain correlation properties, i.e. by a certain relation existing between the successive signal values. Any transmission of the known part of the signal, thus of that part of the amplitude that follows from the correlation properties, is superfluous, for the very reason that it is already known. In any optimum transmission system, therefore, these correlation properties should be discounted at the transmitter and only what is to be regarded as new and individual should be transmitted.

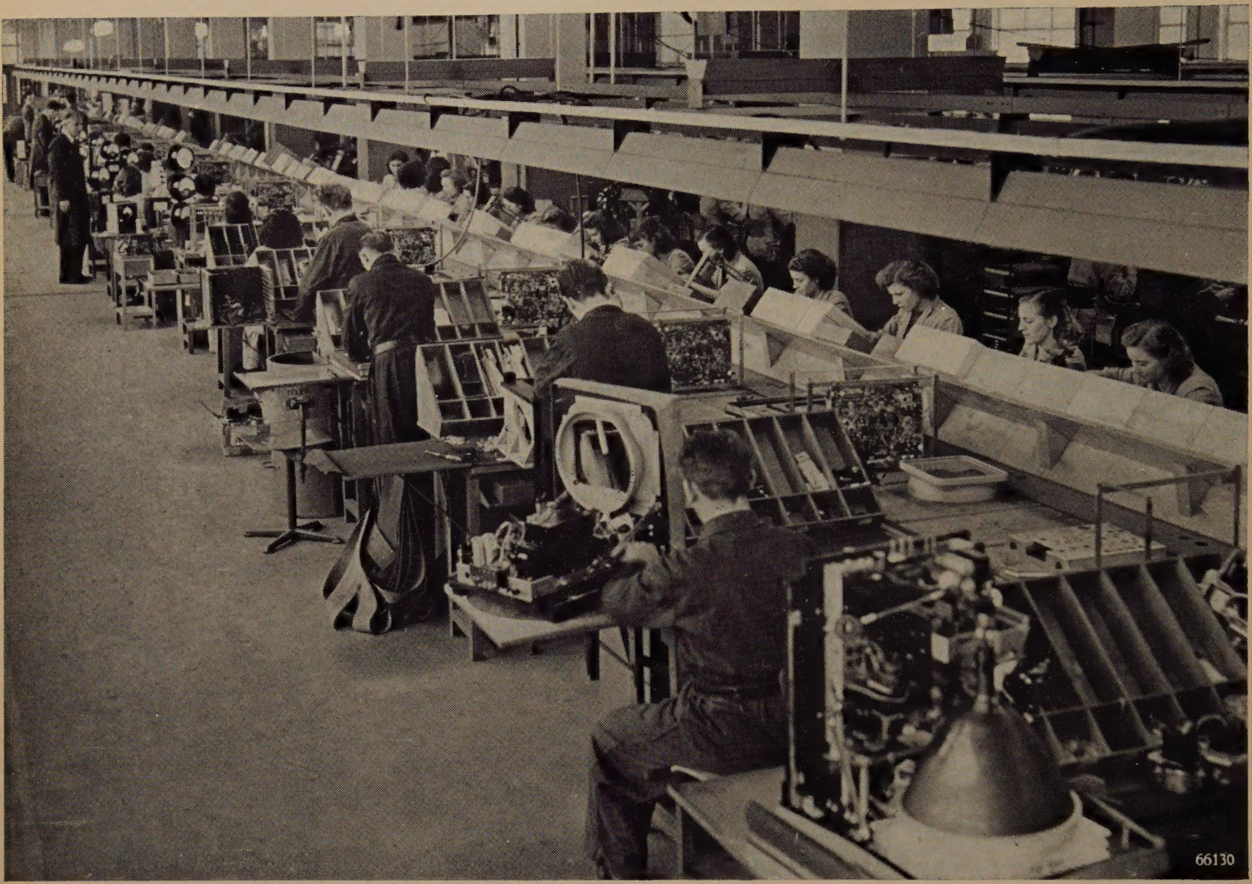
This implies that the signal transmitted in the case of an optimum transmission system possesses the statistical properties of noise, thus having no correlation [4]).

It is the function of the feedback circuit to take into account these correlation properties, which means, therefore, that the properties of the feedback network must be determined by the nature of the signal to be transmitted, on the one hand, and by that of the ear (or eye) on the other hand.

[4]) S. Goldman, Proc. Inst. Radio Engrs **36**, 584, 1948.

Summary. In this article the development of different modulation systems is described, a development which aims at achieving the transmission of audio or video signals as free as possible from interference. The interference occurring in carrier modulation systems and also in pulse-position modulation arises from very small effects of interference being cumulative in the course of transmission. This cumulation can be eliminated by quantizing the signals to be transmitted, both in time and in amplitude. The system of delta modulation, based on this principle, is described and compared with other systems employing quantization (pulse code modulation). It appears that for a good reproduction, using a reasonable bandwidth for the transmitted signal, the apparatus required for delta modulation is relatively very simple, because of coding being applied with the aid of only one unit. This is possible by reason of the fact that with delta modulation account is taken of the correlation properties of the signal, by employing a special kind of inverse feedback in the transmitter, and that the properties of the system are matched with those of speech and music and of the human ear.

THE MANUFACTURE OF TELEVISION RECEIVERS



The first female assembly hand on the right receives a new chassis periodically, mounts on it some parts taken out of trays in front of her, and then passes it on to her neighbour, who in turn mounts on some other parts and passes on the chassis to the next assembly hand, and so on. By the time it reaches the end of the line, on the extreme left of the photograph, the chassis is complete except for the valves and picture tube. It is then placed on a conveyor belt and carried in the opposite direction (in the photograph from left to richt) past a number of points where tests are carried out and any defects traced and corrected. At a certain stage the valves are mounted in their sockets and further tests are carried out. By the time the chassis reaches the assembly hand in the foreground the wooden bottom and the front panel can be put on; here is to be seen also the rubber ring (called a masker) in which, at the next stage, the front end of the picture tube comes to rest. The apparatus in the immediate foreground on the right — the direct-vision receiver type TX 500 U — has the picture tube fitted in.

Further stages, not shown in the photograph, are: testing for picture and sound, running test for a number of hours, completion of the cabinet, and packing of the set.

AUTOMATIC WELDING WITH CONTACT ELECTRODES

by W. P. van den BLINK, H. BIENFAIT and J. A. van BERGEN.

621.791.753.4.002.52

A superficial analysis of a welder's work shows that this depends on two things. The first thing is judgment and experience, enabling the welder to determine how a joint is to be laid, i.e. what kind of joint is to be used, with what type of rod it is to be made, what current is to be used and what the welding speed should be, etc. The second thing is skill in laying the weld — in "holding the arc", keeping up the speed, replacing the rod, "joining up" the new bead, etc. The last-mentioned functions lend themselves to mechanization, which is everywhere desirable for increasing productivity. This mechanization has now become possible through the development of certain types of electrodes together with an automatic welder.

Philips Contact electrodes

The Contact electrodes derive their name from the fact that when welding with a rod of this kind the coating of the rod can be kept continuously in contact with the workpiece, i.e. with these electrodes a joint can be laid by "touch welding".

For a detailed discussion of Contact arc welding and the many aspects of this new method of welding, reference may be made to two articles which appeared in volume 8 of this journal in 1946 [1]). Here we shall only very briefly outline the fundamental principles and the importance of this technique.

It is known that arc welding with normal rods can as a rule be properly carried out only by trained welders. One of the things that a welder has to learn from practical experience is the art of "holding the arc"; the arc between the tip of the rod and the workpiece has to be kept at exactly the same length, so that the welding material from the rod runs smoothly and the penetration of the workpiece is uniform. Moreover, it is of importance to maintain the right length of arc because any deviations, however momentary, may have undesirable consequences. If the arc is too long then the arc voltage rises and in the extreme case the arc may even be extinguished. If, on the other hand, the arc is too short, too little heat is developed and in the extreme case the droplets of metal make a short circuit between the welding electrode and the pool, with the result that the rod may even freeze onto the workpiece, as every welder will have experienced in the beginning of his training.

Thus a very steady hand is required, and it is this which puts a strain on the welder and makes his work tiring, whilst often the results are less consistent than one would wish for.

The metal that is deposited from Contact electrodes is contained not only in the core (wire) but for a part also in the coating in the form of a fine powder. This makes the coating much thicker than that of normal rods, and when welding a very deep cup is formed (*fig. 1*). This cup ensures that the right length of arc is maintained while the rod is kept in touch with the workpiece. In this way it is easy to make a uniform weld and there is no risk of freezing, since the length of arc is constant and much greater than the dimensions of the droplets. Thus with Contact electrodes the arc is maintained as it were automatically.

With normal electrodes, i.e. electrodes having a coating of the normal composition, a deep cup can also be obtained, simply by making the coating very thick, but this would be uneconomical, whilst the

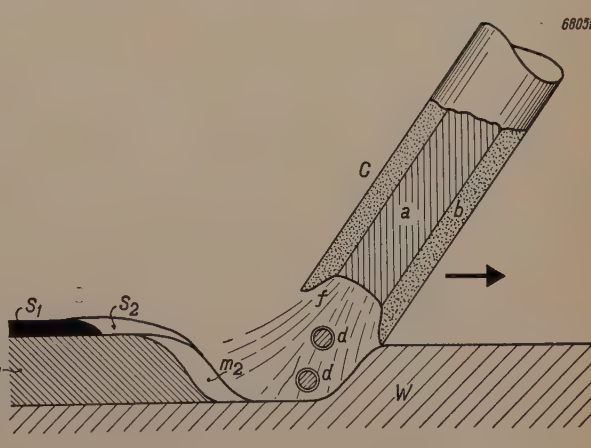


Fig. 1. The melting down of a Contact welding rod. The arrow indicates the direction of travel. C welding rod with core a and coating b (containing iron powder), W workpiece, d droplets of molten iron, m_1 solid and m_2 liquid metal of the weld, s_1 solid and s_2 liquid slag. The depth of the cup f is such as to allow of touch-welding with this electrode.

[1]) P. C. van der Willigen, Contact arc-welding, Philips Techn. Rev. 8, 161-167, 1946; the same author, Penetration and welding speed in contact arc welding, Philips Techn. Rev. 8, 304-309, 1946.

excess of slag-forming material would spoil the quality of the weld. The thickening of the coating of Contact electrodes by incorporating in it a large percentage of the iron that is to be deposited, makes it possible to maintain the optimum ratio of slag to iron. Moreover, there is another advantage of a different nature: Contact electrodes are self-starting. The high iron content makes the coating semi-conducting, so that when it is brought into contact with the workpiece current flows immediately, and owing to the local heating the arc is struck. This dispenses with the troublesome "tapping".

This is not the place to go into the metallurgically favourable properties of Contact electrodes, arising, for instance, from the deep cup protecting the running metal from the atmosphere.

It has just been said that with Contact electrodes the holding of the arc is, as it were, automatically ensured, thanks to the principle of touch welding. All the welder has to do when working with a single rod is to move it along at the right speed for laying the bead. Now it is important to note that the forward travel can also quite easily be made automatic. Furthermore, it is now possible, with the aid of a new apparatus called the Philips Contact automatic welder, for welding to be done entirely automatically with a number of rods in succession. When employing this apparatus welded joints of any length can be made fully automatically.

First we shall deal with the automatic forward travel of the rod and then describe the construction of the automatic welder itself.

Automatic forward travel

In *fig. 2* a sketch is given showing how automatic forward travel can be brought about. This method was actually applied and described some years back by Van der Willigen [2]).

The rod is clamped at a certain angle onto the end of a bar which is set at an angle in the vertical plane and, guided by four insulating rollers, is able to slide down under the force of gravity. Thus the tip of the welding electrode is held against the horizontal workpiece, while the electrode is fed with current via the bar. As the tip of the electrode melts so the bar bearing upon it gradually descends and moves the electrode along parallel to itself. The point of contact between electrode and workpiece, starting at O, thus moves along horizontally according to the rate at which the welding material is deposited until the whole of the rod is consumed. To avoid the arc from being taken over by the clamping device, the bar is, of course, drawn back in good time.

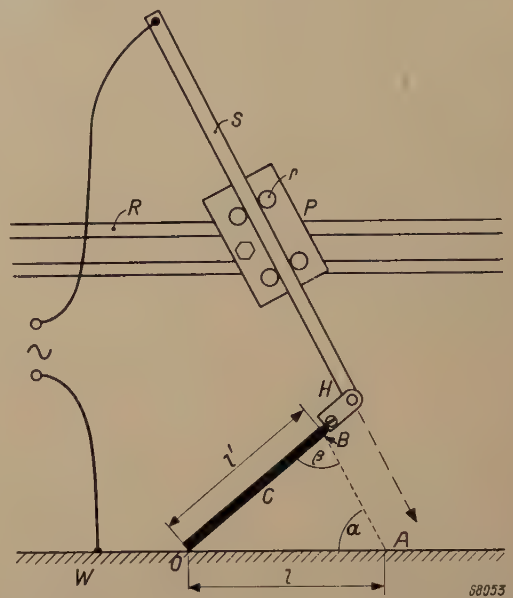


Fig. 2. Set-up for automatic forward travel (for touch-welding). C the Contact welding electrode fixed with adjustable clamp H to the guided bar S. The bar is guided by four insulating, grooved rollers r and by its weight keeps the tip of the electrode pressed against the workpiece W. As the electrode melts down it travels parallel to itself. The mounting plate P, carrying the rollers r, can be fixed to the rail R at any desired angle.

A line drawn through the clamped end of the welding electrode and parallel to the guided bar intersects the workpiece at the point A. $OA = l$ is the length of the bead laid by one rod fully consumed (useful rod length l'). For any welded joint a certain bead length is prescribed, since this determines the thickness of the weld when using a particular kind of electrode (*fig. 3*). The desired bead length can be obtained by a suitable choice of the angles of inclination α and β of the guided bar and the welding

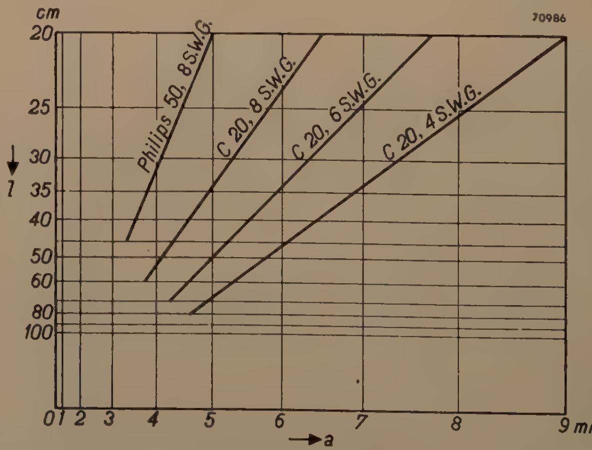


Fig. 3. Diagram for determining the correct bead length l per electrode for a prescribed weld thickness a (in a fillet weld). The lines refer to four different types of welding electrodes.

[2]) P. C. van der Willigen, Contact arc welding, The Welding J., Welding Res. Suppl. 25, 313-320, 1946.

electrode respectively. From fig. 2 it is seen that the condition to be met is:

$$\frac{\sin\beta}{\sin\alpha}=\frac{l}{l'}\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

One of the angles is in the first instance arbitrary. In fact, this is obvious from geometrical considerations: the bead OA to be laid is fixed, but the end B of the welding electrode may lie at any point of the circle with radius l' about the centre O, provided the direction in which the bar is guided is chosen parallel to BA.

In practice it appears that not all combinations of the angles α and β possible according to eq. (1) are equally satisfactory. When for this equation a diagram is plotted from which, with the quotient l/l' and one of the angles given, the other angle can be read, then in that diagram a certain area (i.e. a not sharply defined relation between α and β) can be indicated where the best welding results can be obtained. The position of this area depends to some extent on the type and diameter of the welding electrode, the shape of the weld and the current strength. *Fig. 4* gives such an α-β diagram, in which the optimum area is hatched for a particular case (Contact electrode type C 20, 6 S.W.G. — or C 20-5 — used for a fillet weld made in the flat position, with the maximum permissible current of 350 A). To the left of this hatched area the quality of the weld is adversely affected by the tendency of the slag to run ahead of the arc, whereas to the right of that area the slag has a tendency to lag behind.

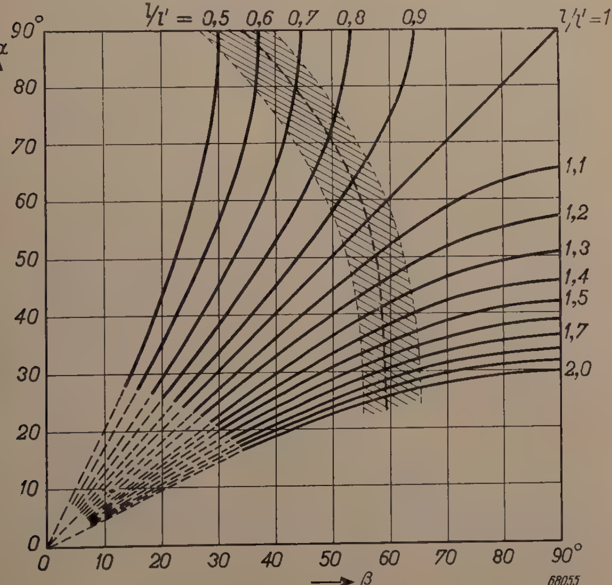


Fig. 4. Diagram for determining the angles of adjustment α and β for automatic welding. The quotient l/l' of bead length and useful electrode length is given, the corresponding curves then showing all combinations of the angles α and β possible for obtaining the desired length of bead. When a fillet weld is to be made in the flat position with electrodes of the C 20, 6 S.W.G. (C 20-5) type, using a current of 350 A, in practice only combinations within the hatched area prove to be satisfactory.

The degree of mechanization of Contact welding achieved by this method of automatic forward travel makes it possible for welding in mass production to be done by less practised welders, or even perhaps by almost unskilled labour, while furthermore the results are more uniform than is usually possible when welding by hand. The fact is that in many cases (depending upon the type of electrode and the current) it takes less than one minute for an electrode to be consumed; the welding speed (rate of travel) is thus very critical if the weld is to be of uniform thickness throughout. With automatic forward travel a constant weld thickness is guaranteed because the rate of travel is determined by the rate of melting of the rod.

Automatic welding of long joints

With the arrangement according to fig. 2 both the holding of the arc and the rate of travel are made automatic, all that is left to be done by hand being the breaking of the arc (drawing back the guided bar) and the starting of a new electrode. We shall now go two steps farther and show how these last two functions can also be carried out automatically.

For laying a long, straight bead requiring a number of welding electrodes, a series of the devices illustrated in fig. 2 can be mounted on an insulated rail set up over the workpiece in line with the joint. The intervals between these automatic welders are chosen equal to the bead length of one rod. Welding rods, e.g. Contact electrodes of the type C 20, 6 S.W.G. or C 20, 4 S.W.G. (continental designation C 20-5 and C 20-6, core diameters 5 mm and 6 mm respectively), are clamped into the holders with their tips resting on the workpiece but insulated from it by a piece of asbestos paper (excepting the first rod). The power supply may provisionally be arranged by connecting each guided bar to the rail with a flexible cable. The rail is connected to one terminal of the supply source and the workpiece to the other. The arrangement is then as sketched in *fig. 5*.

As soon as the current is switched on welding is started with the first electrode (not being insulated by asbestos paper, and the C 20 type of electrodes being self-starting) and the rod is carried along automatically until its clamping device strikes up against the next electrode. As soon as this happens the arc length begins to increase, because

the rod is brought to a standstill while it is still melting. When the arc has been lengthened by a few millimetres (the arc voltage then rising) the atmosphere around the tip of the second electrode will have been sufficiently heated and ionized to

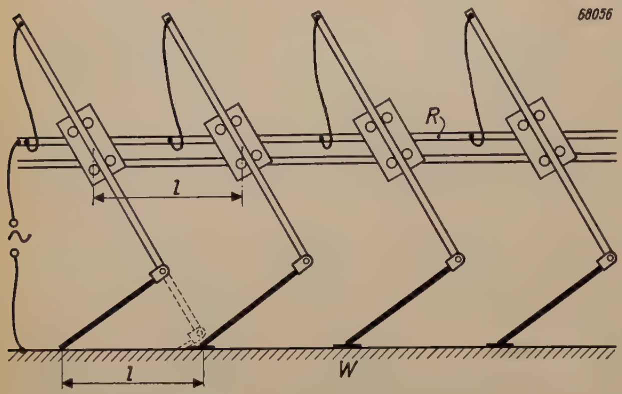


Fig. 5. Diagram of the set-up for laying long welds. A number of welding heads according to fig. 2 are mounted at intervals equal to l on a rail R over the workpiece W. The second, third and following electrodes are temporarily insulated from the workpiece by means of pieces of asbestos paper laid under them. When one electrode has been consumed (dotted position of the guided bar) the next electrode automatically takes over the arc and at that moment the remaining end of the burnt-down electrode has to be drawn up away from the workpiece.

cause the ignition voltage of this electrode to drop below the potential available. Since at the same time the heat melts away the insulating asbestos paper, the second electrode strikes and starts welding [3]). This cycle repeats itself with the whole series of welding rods, so that in this way any length of joint can be welded automatically.

Now we have to consider how the remaining function mentioned above is performed: each electrode holder has to be immediately withdrawn as soon as its electrode has melted away and the next electrode is started. If one holder were not withdrawn in time then the remainder of its electrode would drop further as soon as the next electrode is started, thus making contact with the workpiece again, with the risk of the arc returning from the new electrode to the old one, resulting in an inadequate joint of the new bead and/or even freezing of the clamp to the workpiece. It is therefore highly desirable not to leave the simple function of withdrawing the electrode holder to the care of an operator, but to have this performed automatically too.

[3]) Ordinary paper cannot be used for the insulation instead of asbestos paper because it would burn through too quickly and thus the second electrode would be ignited before the pool from the first one is close enough to it. From the following text it will become clear why this cannot be tolerated.

This can be achieved with the Contact automatic welder, which is a further development of the principle of fig. 2. A photograph of this apparatus is given in *fig. 6*.

The fundamental idea is that the withdrawal of the holder should be brought about by means of the current that begins to flow through the next welding electrode as soon as it is started. One is then sure of each holder being retracted exactly at the right moment, not too soon (before the next electrode has been started) and not too late (when the holder has dropped so far that the arc may return to the old electrode again).

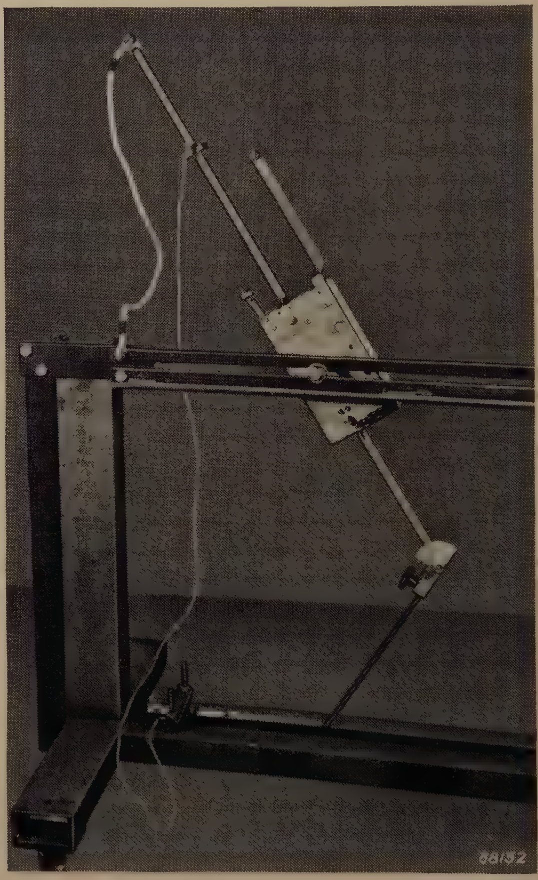


Fig. 6. The Contact automatic welder, mounted in the position for welding.

The Contact automatic welder

The operation of the automatic welder can be described with the aid of the schematic drawing in *fig. 7*, where two "welding heads", connected so as to come into action successively, are mounted on a rail. (Of course more than two welding heads can be mounted in this way.)

The current is not fed to the welding head N direct by the rail but via a terminal c in the preceding head M. As long as the electrode of the head

M is welding, the hinged steel plate *p* is in the position as drawn, being kept there by a leaf spring *b*. In this position the terminal *c* is connected to the rail via the winding of the small electromagnet *e*.

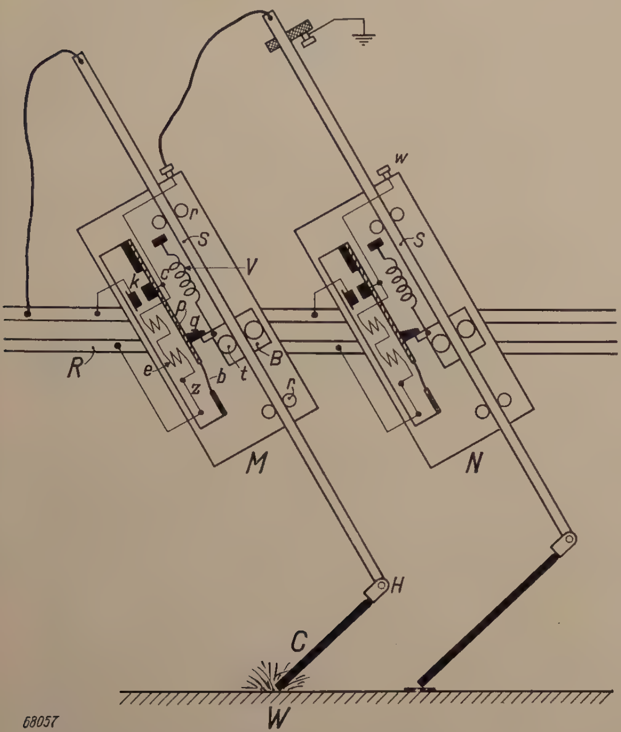


Fig. 7. Schematic drawing of two welding heads *M* and *N*, each fitted with a system for automatically retracting the guided bar. The electromagnet *e* is energized by the current beginning to flow through the welding electrode of the next head. It then attracts the steel armature *p*, the pawl *q* thereby releasing a stop, the spring *V* then clamping and drawing up the bar; at the same time the heavy contact *k* is closed, thereby shorting the winding of the magnet and allowing the welding current, of say 300 A, to flow from the welding head *N* via the terminal *c* and contact *k* to the rail *R*. The armature *p* is held in its new position by the leaf spring *b*, which has two states of equilibrium. *z* is a safety fuse.

As soon as current begins to flow through the welding electrode of the head *N*, this current (at first amounting to only a few amperes) energizes the electromagnet in the head *M*, thereby attracting the armature *p* towards it and releasing a pawl which brings a strong spring into action. This spring first clamps the guided bar onto a block by means of two rollers *t* and then draws up the block with the bar, thus completing the retraction of the remainder of the electrode in the holder of the welding head *M*. At the same time that the armature *p* is drawn towards the electromagnet it closes a heavy contact *k*, thereby shorting the winding of the electromagnet and connecting the terminal *c* — which now has to carry the full welding current of about 300 A for the welding head *N* — direct to the rail. The plate *p* is then again held in its new position by the leaf spring *b*, which has an unstable middle position, thus tending to spring in or out.

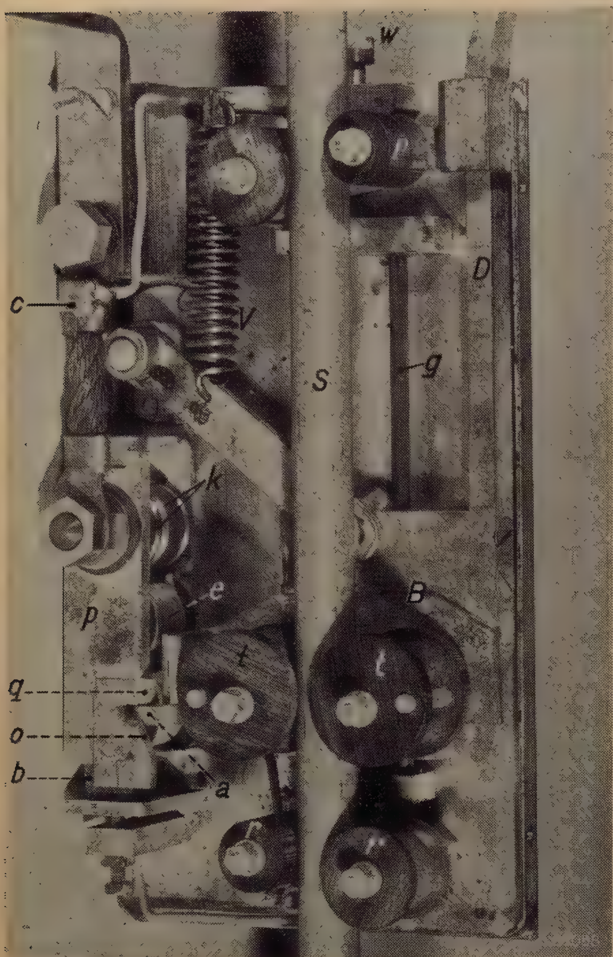
The manner in which the guided bar is clamped and drawn up is illustrated in *fig. 8* and explained in the subscript [4]). In the clamped position the bar can be drawn up still farther for inserting a new electrode in the holder ready for the next welding pass. When the first workpiece has been finished with and the next one is put into position for welding, the spring in each welding head is reset by hand, the guided bar being thereby released and the armature *p* with pawl *q* and the contact *k* returned to its original position. This can likewise be followed in fig. 8.

The terminal *c* in each welding head *N* is connected to a contact *w* on the top of the same head (see figs 6 and 7). Instead of connecting *c* or *w* to the next welding head an earthed contact can be mounted on the bar of the head, the electromagnet of this head then being energized as soon as its own bar has dropped a certain predetermined distance. This system is applied to the last welding head of a series, which thus automatically cuts out and stops the welding of the workpiece.

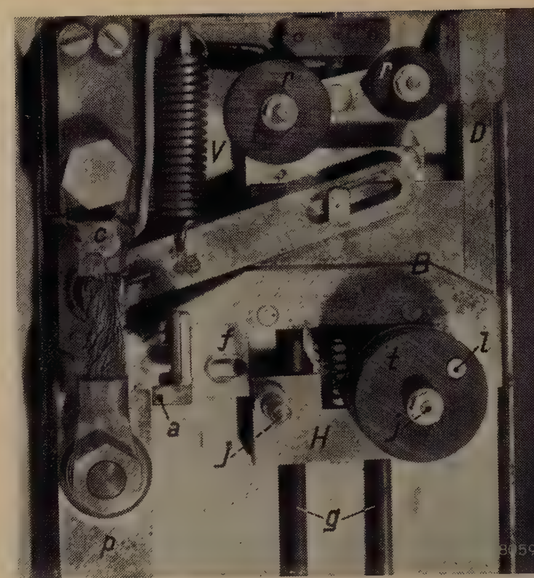
Working with the Contact automatic welder

Obviously the advantages of the automatic welder are most manifest in cases where the same kind of welded joint has to be made over and over again, such as in mass production. The time taken in setting up the rail and positioning the series of welding heads, a job that has only to be done once, is then negligible compared with the time saved in welding. When welding by hand often a great deal of time is wasted in making the joints between one run and the next: when one electrode has been consumed a new one has to be put in the holder, and this often takes so long that in the meantime the pool has set, in which case the slag has to be removed from the end of the last bead, in order to avoid slag inclusions, before starting with the new electrode. When welding automatically with a series of welding heads there is no such interruption, each new electrode taking over the hot bead from the previous one, and when the weld is completed the joints are scarcely perceptible. The useful running time of the welding-current generator is much longer and, as already remarked, the operators do not need so much training, or even practically none at all. Also the operators' time

4) Acknowledgement is due to J. B. van der Wal of Philips Welding Rod Factory at Utrecht (Holland) for his cooperation in designing this construction.



a



b

Fig. 8. *a*) Close-up of the opened welding head, showing, i.a., the parts represented in fig. 7. The sliding block *B* carrying the clamping rollers *t* will glide upward along two pins *g* (only one can be seen here), under the tension of the strong spring *V*, as soon as the stop *a* is released by the pawl *q* being attracted together with the armature *p* by the electromagnet. The block *B* can subsequently be lowered again by pressing down the plunger *D*, thereby stretching the spring *V* again. In this motion of *B*, after sliding over the pawl *q*, the stop *a* forces the plate with pawl, via the sloping face *o*, forward into the starting position, which is stabilized by the leaf spring *b* and in which position the block *B* is then again held by the pawl *q*. *b*) Here the sliding block *B* is in the raised position. The guided bar carrying the welding electrode has been removed, as also one of the two clamping rollers *t*, to show the clamping mechanism. The clamping rollers are rotated about the eccentric spindles *j* fixed in the auxiliary block *H*. When the block *B*, to which the auxiliary block is connected by means of a small spring, is pushed right down then *H* strikes against a stop, the spring is compressed and, owing to the downward movement of the slots *f* in which the pins *l* of the clamping rollers are engaged, the two clamping rollers are rotated a little about *j*, such as to draw them apart. The guided bar is then able to slide freely between the rollers. When, however, the pawl *q* is raised and the block *B* starts moving upward, the small spring forces the auxiliary block *H* away from the block *B*, the two clamping rollers are turned towards each other so as to clamp the guided bar, and as *B* and *H* continue their upward movement the bar is carried along with them.

In this position the guided bar can be drawn up further by hand if desired, the frictional force acting upon the clamping rollers then rotating them away from each other and thus reducing the clamping force. The guided bar cannot, however, slide down because then the frictional force acting upon the clamping rollers would cause the latter to rotate towards each other and thus increase the clamping force.

can be fully utilized, since the stubs of the electrodes can be removed and new electrodes put in the holders for the next run while the welding is proceeding. If sufficient power sources are available it is possible for two series of welding heads to be operated simultaneously by one man, either on two workpieces or on the same workpiece if it is of great length. The latter method has the advantage that excessive transverse shrinking can be avoided: the two series can be started from the same point in the middle of the workpiece and continued in opposite directions outward.

The power-supply rail can be set up according to the nature of the workpieces to be welded. If they are easily transported a permanent set-up is recommended, with a specially adapted bench or jig for quickly aligning the workpieces. If the workpieces are so heavy that they can only be lowered onto the bench with an overhead crane then it may be advantageous to be able to swing the rail back by hingeing it onto a wall or frame. In cases where the workpiece cannot be transported, as for instance when welding a floor or a ship's deck, a mobile frame carrying the rail and the welding heads is necessary.

Fig. 9 shows an example of welding a workpiece in a fixed position.

In principle overhead or vertical welding is also possible, with welding heads of a somewhat modified form. The pressure then required on the electrodes can be applied by means of a set of weights with cables passing over pulleys.

The foregoing should suffice to show that the Contact automatic welder is a very flexible machine easily adapted on the one hand to the welding equipment available and, on the other hand, to the requirements of the particular job.

Summary. Since Contact welding electrodes allow of touch-welding, the welder is relieved of the task of "holding the arc", this being done, as it were, automatically. Also the forward travel of the rod and the change-over from one electrode to the next can be carried out automatically, so that welds of any length can be made entirely automatically. For this purpose Philips have developed the Contact automatic welder. This consists of a "welding head" fitted with an oblique bar sliding in its own longitudinal direction and with the welding electrode clamped to its lower end at a given angle. The weight of the bar presses the tip of the electrode onto the workpiece and as the electrode melts down it is carried along parallel to itself. The rate of travel (and thus the thickness of the weld) is adjusted by varying the angles at which the bar and the electrode are set. A series of such welding heads are mounted at appropriate intervals along a rail above the workpiece, and each electrode comes to rest at the point where its bead begins. Each electrode takes over the arc from the previous one automatically as soon as the latter is consumed, the preceding electrode holder then being automatically retracted by means of an electromagnet in the head. In this way welding can be done in mass production with unskilled personnel, while still obtaining very good and uniform welds. The joints between the beads are scarcely perceptible.

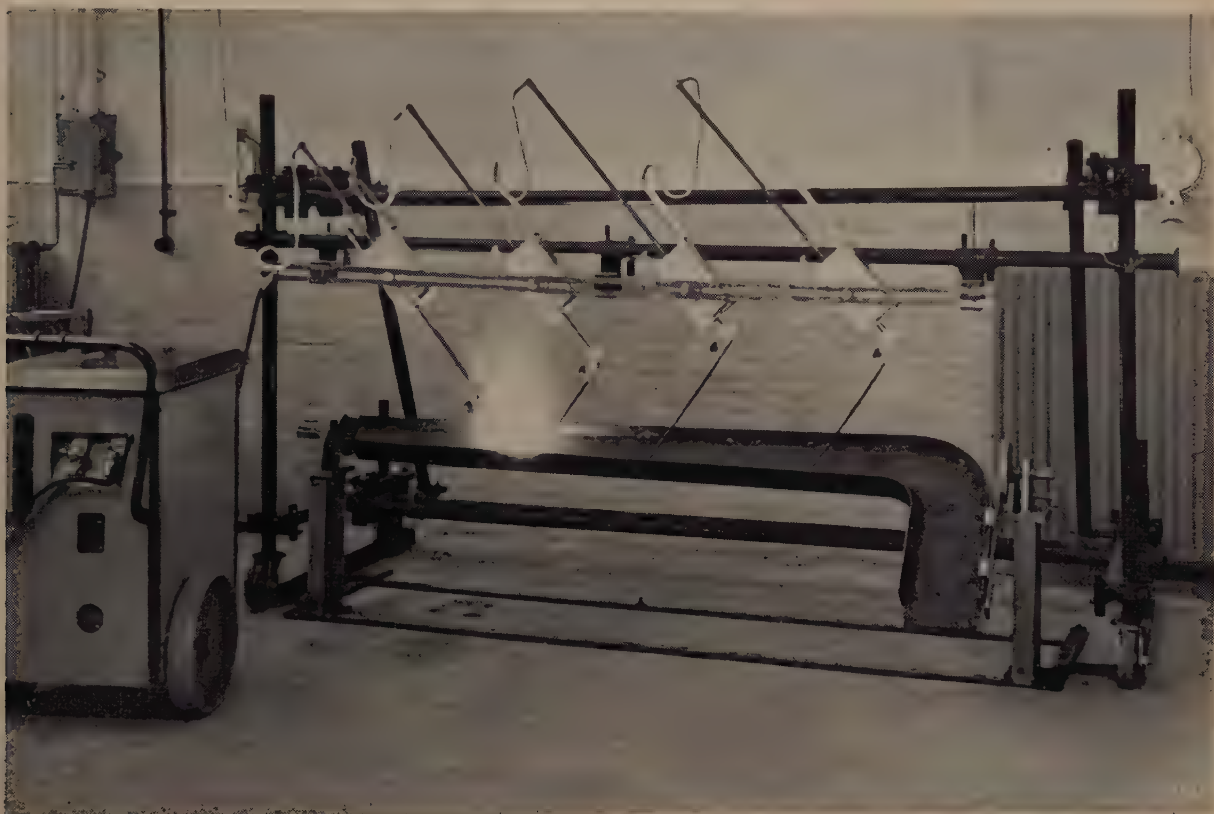


Fig. 9. The Contact automatic welder at work. The electrode of the first welding head has been consumed and the bar drawn up. Welding is being continued with the second welding head.

APPARATUS FOR TESTING TRANSISTORS

by P. J. W. JOCHEMS and F. H. STIELTJES.

621.314.632:621.315.59:
621.317.7

In America some years ago a circuit element was invented which, like the triode, is able to amplify a signal. In its most common form this circuit element, called transistor, consists of a block of germanium with one large electrode and two small, pointed electrodes placed in close proximity to each other.

The transistor is still in a stage of development, the ultimate outcome of which is uncertain. The subject, however, is of such importance that Philips have also taken it up energetically. The testing apparatus described here has proved to be very useful in carrying out these investigations.

In the early stage of radio-telegraphy the crystal detector was for many years the most reliable and most used detector, until in the first world war it had to make place for the triode, which has the advantage of being able to amplify a signal at the same time. Some twenty years later, with the advent of centrimetric-wave technique (radar), for which at first there were no suitable tubes available, the crystal detector again came to the fore.

There has been no lack of attempts to develop this detector into a circuit element capable of functioning as an amplifier, like the triode. In America these efforts have been crowned with success by the invention of the transistor [1]). It is not the intention of this article to give a detailed account of the working and uses of the transistor, which can be found in the literature on the subject [2]). Here only a brief summary will be given, before proceeding to describe two types of apparatus with which the main characteristics and the gain factor, respectively, of transistors can be determined.

Principle of the transistor

In its most common form a transistor consists of a semi-conducting crystal, soldered onto a metal plate (the base, B in *fig. 1*), and two metal points bearing resiliently on the crystal. A suitable material for the semi-conducting crystal is germanium with a surplus of conduction electrons ("type N"). One of the metal points is called the emitter (E), because it sends into the crystal a controlling current, and with it charge carriers, while the other is called the collector (C), because it collects the charge carriers introduced with the controlling current.

The contact of each of the metal points with the germanium has a rectifying action, and in the case of type N germanium this allows a positive current to flow much more readily from the point to the germanium than in the other direction.

The working of the transistor can be demonstrated with the aid of the circuit represented in fig. 1. A small potential, positive with respect to the base, is applied to the emitter, such that a current of 0.5 to 1 mA flows in the forward direction. The collector is given a much higher negative potential, such that a current of several milliamps flows in the opposite direction. Connected in series with the collector is a resistance, the load resistor R_2, of say 20,000 ohms (i.e. much higher than the resistance in the emitter circuit).

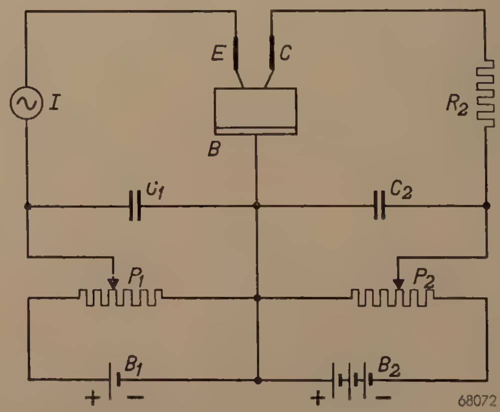


Fig. 1. Amplifying circuit with a transistor (germanium crystal with base B, emitter E and collector C). I source of input current or voltage. R_2 load resistor. B_1 and B_2 direct-voltage sources with potentiometers P_1 and P_2 and bypass capacitors C_1 and C_2.

When the emitter current is slightly varied, e.g. by incorporating an alternating-current source in the emitter circuit, then a variation in current, i.e. an alternating current, occurs likewise in the collector circuit. This is particularly the case when

[1]) J. Bardeen and W. H. Brattain, The transistor, a semi-conductor triode, Phys. Rev. **74**, 230-231, 1948.

[2]) See, e.g., J. Bardeen and W. H. Brattain, Physical principles involved in transistor action, Phys. Rev. **75**, 1208-1225, 1949, and W. Shockley, Electrons and holes in semi-conductors, with applications to transistor electronics, Van Nostrand Co., Inc., New York 1950.

the emitter and the collector are very close together, say at a distance of 0.05 mm. It is even possible for the amplitude of the alternating current in the collector circuit to be greater than that in the emitter circuit. But even without that there may be a considerable A.C. power gain, since the resistance in the collector circuit is much greater than that in the emitter circuit.

Thus, under suitably chosen conditions, the transistor is able to amplify a signal, and when the output is fed back to the input it can be caused to oscillate. So far, it is therefore like the triode. Practical advantages of the transistor over the triode (or other amplifying valves) are: (1) the absence of a heated cathode, (2) the smaller dimensions in which it can be made (*fig. 2*), and (3) its practically

Fig. 2. True-to-scale photograph of a Philips transistor, with an EF 42 pentode for comparison. (It should not be concluded that the transistor and the pentode could replace each other.)

unlimited life under normal use. These advantages make the transistor particularly attractive for those applications where hitherto large numbers of tubes have been employed, e.g. in telephone repeaters and electronic computers. Before that stage is reached, however, there are still some difficulties to be overcome.

Notwithstanding the analogy mentioned between a transistor and a triode, there is an important difference in their working. The triode — provided it is so adjusted that no grid current flows and if it is used at not too high frequencies — does not take up any input power. The input circuit (emitter circuit) of the transistor, on the other hand, has a fairly low impedance and a certain amount of power is absorbed. The consequences of this for the transistor technique are: (1) that it is just as important to speak of power gain as of voltage gain and (2) that account has to be taken of the matching, that is to say, that the internal resistance of the input-voltage source and the load resistance each have in themselves the most favourable value to give either maximum power gain or undistorted output, as required. (As a matter of fact the same applies for valves which are used in such a way that they do need a certain input power, as for instance transmitting tubes, which are mostly so adjusted that grid current flows.)

In what follows, an apparatus is described by means of which the most important characteristics of a transistor can be displayed on the screen of a cathode-ray oscilloscope and the power gain can be measured directly.

Oscillographic recording of transistor characteristics

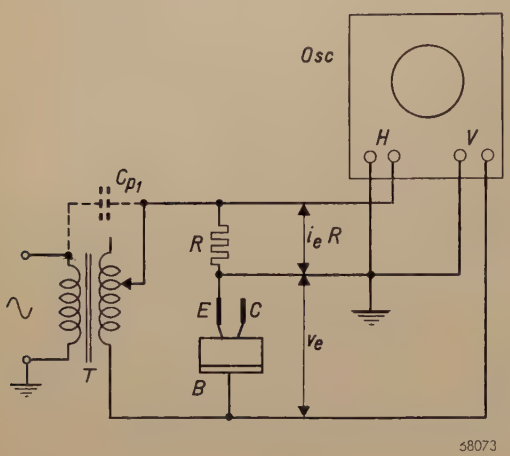
When dealing with a transistor one has to consider the currents i_e and i_c flowing through the emitter and the collector respectively, and the voltages v_e and v_c between the emitter and collector, respectively, and the base. (With all these symbols the instantaneous values are meant.) Each of these quantities plotted as a function of one of the others forms a transistor characteristic.

In principle, for instance, the emitter characteristic $v_e = f(i_e)$ at $i_c = 0$ can be traced on an oscilloscope with the aid of a circuit as represented in *fig. 3a*. A variable alternating voltage derived from a transformer is applied in series with the transistor (emitter and base) and a resistor R. The voltage across R is proportional to the current i_e flowing through the transistor. This voltage is applied to the input terminals for the horizontal deflection of a cathode-ray oscilloscope, the voltage v_e (between the emitter and the base) being applied to the input terminals for the vertical deflection. An oscillogram is then obtained of the characteristic $v_e = f(i_e)$.

When, however, a normal oscilloscope is used for this purpose one is faced with two difficulties. The first is that the voltages v_e and i_eR are too small to give sufficient deflection without being amplified. Also owing to the rectifying action both these voltages have a direct-voltage component. The amplifiers of a normal oscilloscope such as the type

GM 3159 [3]) do not amplify direct voltages, so that this component cannot be represented in the oscillogram. Consequently the position of the coordinate axes — the direction of which is known — remains undetermined. Other types of oscilloscopes, such as the GM 3152 [4]) and the GM 3156 [5]), have only one amplifier, and this again is not a direct-voltage amplifier.

is therefore no longer proportional to the current i_e at every instant, since only part of the alternating-current components of i_e flow through R, and moreover there is a stray alternating current flowing from the non-earthed mains terminal via C_{p1} and R to the other (earthed) terminal. As a consequence of all this, the forward and return lines of the characteristic in the oscillogram do not exactly



a

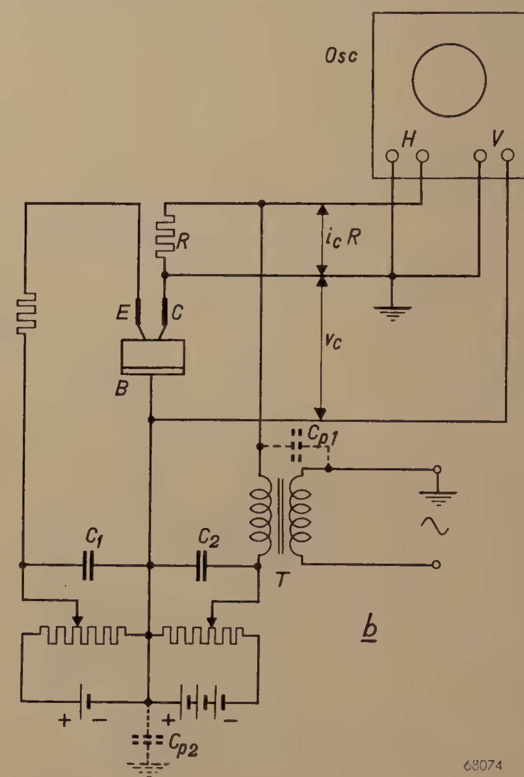


Fig. 3. Elementary circuits for tracing (*a*) the emitter characteristic with collector current zero, and (*b*) the collector characteristic with constant emitter current. *B-E-C* transistor (cf. fig. 1). *T* transformer. *R* series resistor. *Osc* cathode-ray oscilloscope with two amplifiers; *H* input terminals for the horizontal, *V* for the vertical deflection. Owing to the direct-current components superimposed on the transistor currents as a consequence of the rectifying action, the two amplifiers would have to be direct-voltage amplifiers. The stray capacitances (C_{p1}, C_{p2}) arising from the one-sided earthing of the input terminals cause the formation of a loop in the oscillogram (see fig. 4).

The second difficulty arises from the fact that in a normal oscilloscope the inputs of the two amplifiers are asymmetric (earthed on one side) and thus have one common terminal.

In the first place this fact precludes the use of such amplifiers in cases where there is *no* common terminal for the two pairs of terminals whose voltage differences are to be traced as functions of each other. But even when the pairs of terminals have one point in common (in the case of fig. 3*a* the emitter), which is then earthed, interference arises as a consequence of inevitable stray capacitances. The capacitance C_{p1} between the coils of the transformer in fig. 3*a*, for instance, is in parallel with the resistance R. The voltage across R

coincide but form quite a wide loop (see *fig.* 4).

Further, the capacitance of the D.C. sources with respect to earth may form a shunt, as seen from fig. 3*b*. The object of this circuit is to record the collector voltage v_c as a function of the collector current i_c for given values of the emitter current. If, owing to the oscilloscope being connected to the circuit, the collector is earthed, then the stray capacitance C_{p2} of the D.C. sources comes to lie between the collector and the base, whilst the stray capacitance C_{p1} of the transformer T is shunted across the resistance R, from which the voltage for the horizontal deflection is derived.

When the characteristic of a vacuum diode is to be traced it is already known in advance that at low frequencies (where the effects of electrode capacitance and transit time of the electrons are negligible) a properly recorded characteristic will not

[3]) Described in Philips Techn. Rev. **9**, 202-210, 1947.
[4]) Described in Philips Techn. Rev. **4**, 198-204, 1939.
[5]) Described in Philips Techn. Rev. **5**, 277-285, 1940.

show any loop. So if a loop does appear in the oscillogram this can be ascribed entirely to a fault in the measuring equipment, as indicated in the case of fig. 3. In such a case a small loop is not very troublesome. With a new element such as the transistor,

Fig. 4. Emitter characteristic traced by the system according to fig. 3*a*. *1* forward direction, *2* inverse direction (with perceptible looping). There are no coordinate axes.

however, one is not so sure that the formation of loops is due entirely to the apparatus employed. In the case of a breakdown, for instance, — we shall see an oscillogram of this presently — there is a noticeable development of heat which, owing to thermal inertia, may result in a real loop. For these purposes, therefore, the measuring equipment should be quite free of the fault in question.

An oscilloscope quite suitable for the purpose, which has two symmetrical direct-voltage amplifiers, is at the moment still in a stage of development. But a satisfactory solution has been reached by employing existing Philips' equipment, as will be shown in what follows.

Tracing the coordinate axes

The first objection — the absence of coordinate axes — can be met in a simple way by replacing each of the amplifiers of the oscilloscope by an electronic switch, one input of which is not used. The circuit is then as represented diagrammatically in *fig. 5*, where for the sake of clarity the electronic switches have been drawn as ordinary two-pole two-way switches, thus ignoring the fact that they act as amplifiers. This manner of representation — which implies that a direct-voltage component of the input voltages is also transmitted — is permissible because owing to the switching these direct voltages are converted into alternating voltages, of rectangular wave form, which, provided the switching frequency is sufficiently high, are transmitted without any distortion.

The switches S_V and S_H are to be imagined as operating at different frequencies. These frequencies are chosen, for instance, higher than the frequency of the voltage supplied by the transformer T. There are then always intervals where both switches are in the position shown in fig. 5 with fully drawn lines. A part of the characteristic — in the present case $v_e = f(i_e)$ — is then displayed on the screen of the oscilloscope. A moment later S_V, for instance, will have assumed the position indicated by a broken lines. The vertical-deflection voltage is then zero, but not so the horizontal-deflection voltage, so that the spot describes on the screen a small part of the horizontal axis. A little later both switches are in the broken-line positions and the spot is then at the origin. Finally, in the intervals during which S_V is in the fully-drawn position and S_H in the broken-line position a part of the vertical axis is traced on the screen. Thus the oscillogram shows the characteristic (as a broken line), the two coordinate axes (at least for the part corresponding to the sweep of the currents or voltages) and the origin.

An electronic switch suitable for our purpose is the type GM 4580 [6]), the switching frequency of which is variable between 2 c/s and 40,000 c/s (with the latest type, GM 4580/01, up to 50,000 c/s).

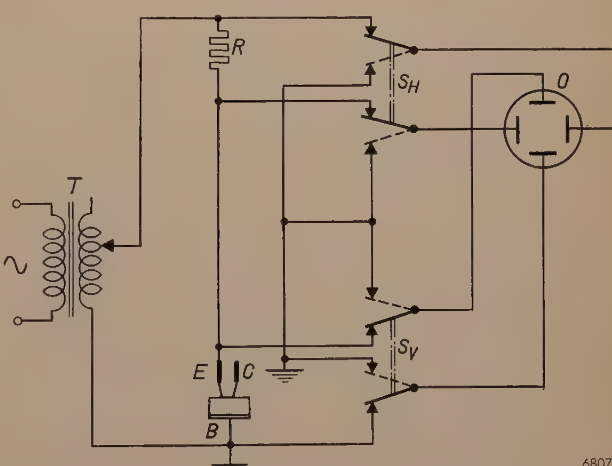


Fig. 5. The switches S_H and S_V represent electronic switches taking the place of the amplifiers in the oscilloscope *Osc* of fig. 3. *O* oscilloscope tube. With this system not only the characteristic but also the two coordinate axes and the origin are traced.

Tracing two voltages without a common terminal

When the inputs of the amplifiers of the oscilloscope have one common terminal, and as a consequence the point of the circuit connected to that terminal is earthed, then, as shown with the aid

---

[6]) Described in Philips Techn. Rev. **9**, 340-346, 1948.

of fig. 3, stray capacitances have a disturbing influence. Two entirely independent inputs, on the other hand, permit of some point being earthed which will eliminate this influence. If, for instance, the oscilloscope of figs 3*a* and *b* had this property, the point B could be earthed, in which case the stray alternating current originating from the mains and flowing via C_{p1} would no longer flow to earth through R but through the transformer secondary, and then neither C_{p1} would form a shunt across R nor (in fig. 3*b*) would C_{p2} form a shunt across the collector. Moreover, with an oscilloscope having independent inputs it would, of course, be possible to trace the voltage v_V between two terminals *1* and *2* as a function of the voltage v_H between two other terminals *3* and *4*, independently of any voltage between terminal *1* or *2* on the one hand and terminal *3* or *4* on the other hand.

It will be shown that this independency of the two inputs can be obtained when the amplifiers (in this case the electronic switches) have a symmetrical input and are provided with negative feedback in a particular way, as represented in *fig. 6*.

Let us consider first only one of the amplifiers, say A_V. The potentials of the input terminals with respect to earth are denoted by v_a and v_b, those of the output terminals by v_A and v_B. Now:

$$v_a = \tfrac{1}{2}(v_a + v_b) + \tfrac{1}{2}(v_a - v_b)$$

and

$$v_b = \tfrac{1}{2}(v_a + v_b) - \tfrac{1}{2}(v_a - v_b)\,,$$

where $v_a - v_b$ is the input voltage v_V and $\frac{1}{2}(v_a+v_b)$ is the mean potential of the input terminals with respect to earth.

The gain m, which we shall assume to be much larger than unity, is the ratio of the voltage difference $v_A - v_B$ between the output terminals to the voltage difference $v_a - v_b$ between the input terminals. Thus:

$$v_A - v_B = m(v_a - v_b) = mv_V\,.$$

The quantity $v_A - v_B$ determines the magnitude of the vertical deflection on the oscillograms. Thus this deflection — except for a condition to be dealt with presently — is independent of the voltage $\frac{1}{2}(v_a+v_b)$. This means that v_a and v_b need not be regarded as voltages with respect to earth, but that they may equally well be regarded as voltages with respect to some other point having itself an arbitrary potential with respect to earth.

This applies, of course, also to the horizontal deflection if the respective amplifier (A_H, input voltage v_H) is constructed in the same way. It is therefore possible to trace a voltage v_V as a function of a voltage v_H independently of the voltage level of each pair of input terminals, thus without it being necessary for the terminal *1* or *2* to have the same potential as the terminal *3* or *4*.

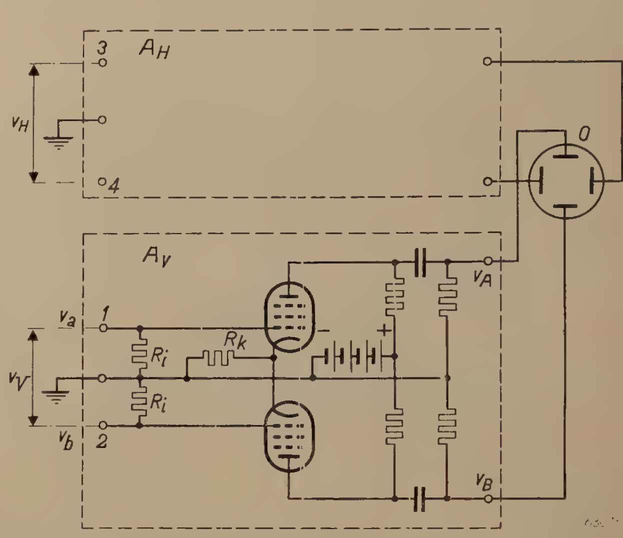


Fig. 6. Oscilloscope tube O with identical amplifiers A_H and A_V, for the two directions of deflection. R_i input resistors, R_k cathode resistor. The oscillogram obtained represents $v_V = f(v_H)$, to a certain extent independently of the voltage across the terminal pairs *1-2* and *3-4*.

So far, however, no account has been taken of the influence that the sum voltage, v_a+v_b, may have not only upon the amplitude of the deflection but also upon the focusing of the oscillogram.

The first effect of this is that $\frac{1}{2}(v_a+v_b)$ might be so large as to cause the valves to be overloaded, distortion then occurring in the amplifier. This is amply counteracted by connecting a large (non-bypassed) resistor R_k in the common cathode lead of the valves connected in push-pull. The sum voltage influences the two cathode currents in the same sense, so that the current through R_k changes according to the sum voltage and a large proportion of $\frac{1}{2}(v_a+v_b)$ is taken up in R_k. In other words, R_k gives a heavy negative feedback for the sum voltage.

The difference voltage, $v_a - v_b$, on the other hand influences the two cathode currents in opposite senses and thus the current passing through R_k remains unchanged. The gain m of the difference voltage therefore remains unaffected by the introduction of R_k.

The voltages from the output terminals are now

$$v^A = n\cdot\tfrac{1}{2}(v_a + v_b) + m\cdot\tfrac{1}{2}(v_a - v_b), \quad . \quad . \quad (1)$$

and

$$v_B = n\cdot\tfrac{1}{2}(v_a + v_b) - m\cdot\tfrac{1}{2}(v_a - v_b), \quad . \quad . \quad (2)$$

where $n \ll m$.

As is the case without the feedback resistor, the voltage $v_A—v_B$ determining the amplitude of the deflection is thus equal to $m(v_a—v_b)$.

However, importance is also to be attached to $\frac{1}{2}(v_A+v_B)$, the mean potential of the output terminals with respect to earth, since this voltage is at the same time the mean potential of the pair of deflecting plates with respect to the (earthed) accelerating anode of the cathode-ray tube. This potential defocuses the electron beam and therefore, if a sharply focused oscillogram is to be obtained, it has to be kept low. This defocusing is the second effect of v_a+v_b referred to above.

From (1) and (2) it is found that

$$\tfrac{1}{2}(v_A + v_B) = n\cdot\tfrac{1}{2}(v_a + v_b)\,,$$

and from this it is seen that in order to minimize the effect under consideration n has to be made small, which is a further reason for having a heavy negative feedback for the sum voltage.

Further particulars about the use of the circuit

The circuit at which we have arrived in the foregoing is that depicted in *fig. 7*. It consists of two electronic switches GM 4580 (H for the horizontal, V for the vertical deflection) and an oscilloscope tube with accessories. Owing to the symmetrical output of the electronic switches a tube is used which has four symmetrical deflecting plates, e.g. the type DN 9-4.

The most important parts of one of the electronic switches are shown in fig. 7. These comprise two double pentodes EFF 51, each with a cathode resistor R_{kI} and R_{kII}. An auxiliary current, i_{comm}, of rectangular wave form, derived from a multivibrator (not shown here), is sent through each of these resistors. These two auxiliary currents (switching currents) are in antiphase, so that the tubes are alternately cut off by the voltage $i_{\mathrm{comm}}R_k$. The frequency of the current i_{comm} is variable, as already mentioned.

A variable positive or negative direct voltage (E_I, E_{II}) can be applied to one of the control grids of the EFF 51 tubes to shift the oscillograms corresponding to the inputs I and II with respect to each other [7]), in the horizontal direction in the case of the upper electronic switch in fig. 7 and in the vertical direction in the case of the lower switch. These direct voltages do not as a rule interfere with the circuit to which the electronic switch is connected, but such cannot be said to be the case when the object under test is a transistor. From fig. 7 it is seen that E_{I_V} for instance affects the D.C. adjustment of the transistor (though via high resistances). For this reason the variable direct voltage at the two inputs used (I in fig. 7) has been made zero (this is the only change that had to be made in the electronic switches GM 4580). At the other two inputs, which are not connected to any other point, use is then made of the variable direct voltage (E_{II}) in order to get the coordinate axes in the right place. This is brought about in the following way.

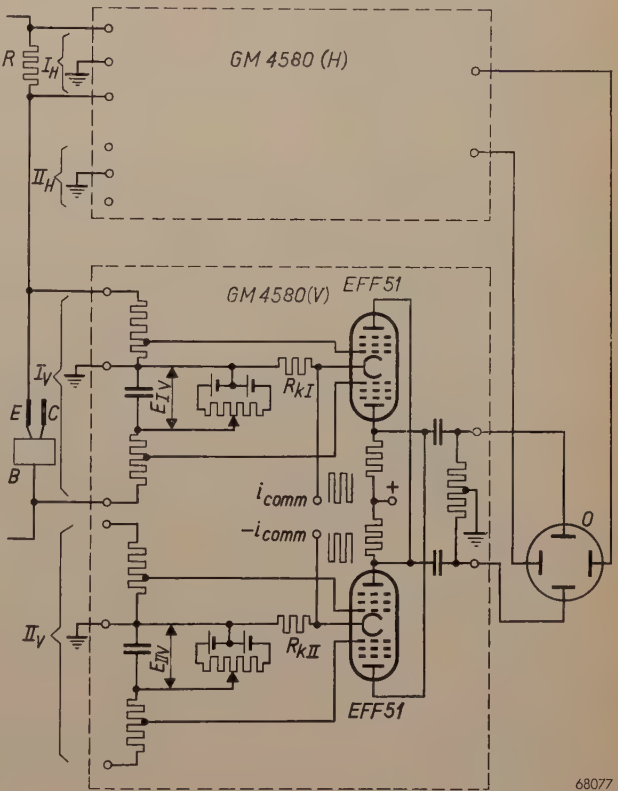


Fig. 7. Oscilloscope tube O with two electronic switches GM 4580 (H = horizontal, V = vertical), each with two symmetrical inputs and two double pentodes EFF 51. R_{kI}, R_{kII} cathode resistors. i_{comm} and $-i_{\mathrm{comm}}$ switching currents. E_{I_V} and E_{II_V} variable direct voltages (here the voltages E_{I_V} and E_{I_H}, not indicated, have to be zero).

The apparatus is started up, but without sending current through the transistor. Owing to the switching each of the electronic switches then supplies an alternating voltage of rectangular wave form, the amplitude of which depends upon E_{II_V} or E_{II_H} respectively. These voltages are not of the same frequency, since the switching frequencies have been differently adjusted. As is easily understood, the oscillogram then shows four dots situated in the corners of a rectangle. The width of the

[7]) See the article quoted in note [6]), pp 345 and 346.

rectangle depends upon the direct voltage E_{II_H} at the electronic switch for the horizontal deflection, while its height is governed by the direct voltage E_{II_V} at the swith for the vertical deflection. By changing these direct voltages the dots can be made to coincide with each other, this then giving the correct adjustment. The fact that for this adjustment E_{II_H} and E_{II_V} are not as a rule zero is due to small differences between the valves. In course of time these differences may vary slightly, so that the adjustment has to be repeated now and again.

Upon a current being sent through the transistor both its characteristic and the two axes appear on the screen, as already described.

An oscilloscope tube is rather more sensitive at the pair of deflecting plates nearest to the electron gun than at the other pair. It is therefore advisable so to adjust the gains of the electronic switches that this difference is just compensated. In what follows it will be assumed that this has been done.

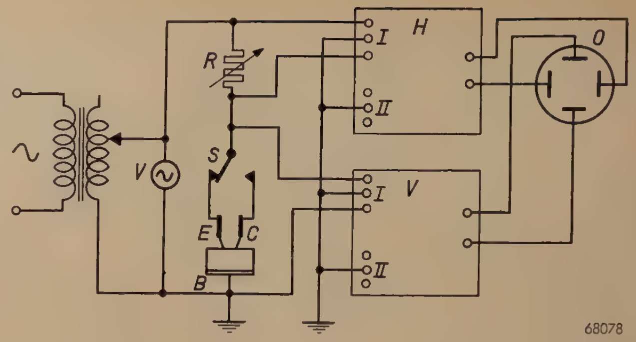


Fig. 8. Circuit for tracing the emitter and collector characteristics.

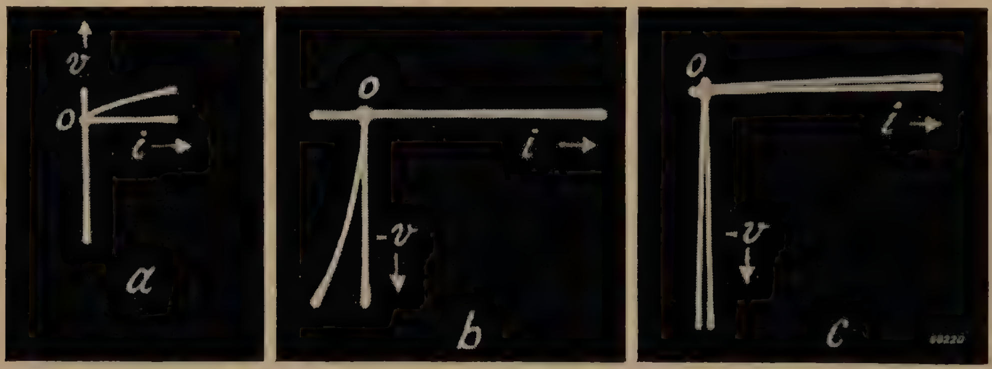


Fig. 9. Diode characteristics of a transistor with different values of the load resistor R: (a) R of the same order as the resistance in the forward direction, (b) R of the same order as the resistance in the inverse direction, (c) intermediate case.

Some transistor characteristics

Emitter characteristic at $i_c = 0$ and collector characteristic at $i_e = 0$

Fig. 8 represents the circuit with which, according to the position of the switch S, one of the two so-called diode characteristics (the emitter characteristic for collector current zero and the collector characteristic for emitter current zero) can be traced.

As already remarked, both the emitter and the collector contacts have a rectifying action. Thus the diode characteristics show a sharp bend near the origin. It depends upon the value of the resistor R which of the two parts of the characteristic shows up clearly: if the value of R is of the same order as the resistance in the forward direction (*fig. 9a*) then the forward characteristic is clearly shown and the characteristic of the inverse direction almost

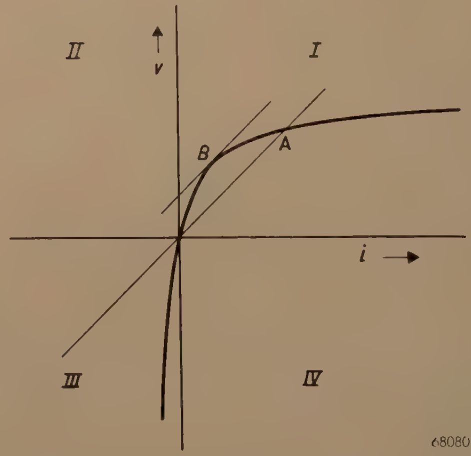


Fig. 10. Diode characteristic of a transistor on an enlarged scale. In the quadrant I lies the characteristic of the forward direction, in quadrant III that of the inverse direction. At the point A where the characteristic is intersected by a line drawn through the origin at an angle of 45° to the axes the resistance of the emitter (or collector) is equal to the value of the load resistor R. At B, where the tangent to the characteristic makes an angle of 45° with the axes, the difference resistance is equal to R.

coincides with the negative v-axis; if the value of R is much greater (fig. 9*b*) then the forward characteristic lies approximately on the i-axis, while the inverse characteristic is clearly seen. Fig. 9*c* represents an intermediate case where both parts of the characteristic can still be distinguished. It is for this reason that the resistor R has been made continuously variable, from 1800 ohms to 200 000 ohms. The value of R can be read from a scale.

In *fig. 10* the characteristic is shown on an enlarged scale. A line drawn through the origin at an angle of 45° to the axes intersects the characteristic at a point A, where the ratio of voltage and current is just equal to R. At the point B, where the tangent is at an angle of 45°, the difference resistance is equal to R. If it is desired to know at what value of current and voltage the resistance, respectively the difference resistance, has a certain value R', then the characteristic has to be traced with $R = R'$ and the points A and B determined along the curve.

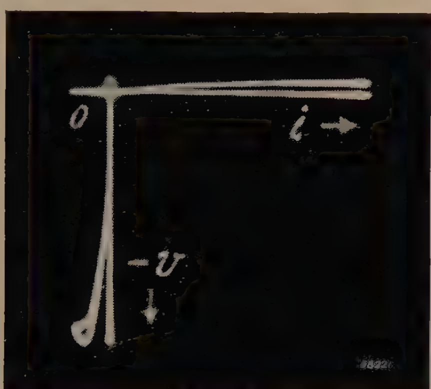


Fig. 11. When the voltage is raised too high breakdown occurs, this being made visible in the oscillogram because when the current is raised beyond a certain value the voltage drops. When measurements are taken with not too low frequencies (say 50 c/s) a loop is formed in the oscillogram.

A suitable value of the alternating voltage for carrying out the test is 100 V (r.m.s. value). Upon the voltage being raised a value is ultimately reached where breakdown occurs. This is clearly seen on the oscillogram (*fig. 11*); when the current is raised beyond a certain level this causes the voltage to drop. A loop is then formed (this time a real one!). The voltage at which this takes place can be read from a voltmeter.

Collector characteristics with the emitter current as parameter

When the collector voltage is traced as a function of the collector current for different values of the emitter current (kept constant in each case), a family of curves is obtained as represented in *fig. 12*.

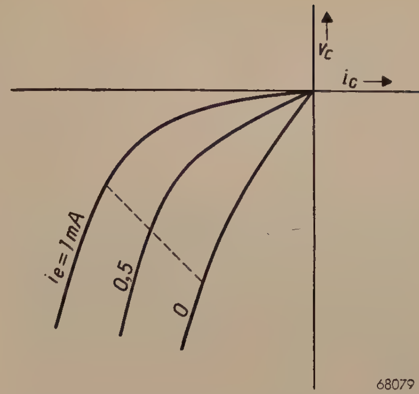


Fig. 12. Collector characteristic $v_c = f(i_c)$ at different constant values of the emitter current i_e.

When the emitter current is varied, say from 0 to 1 mA, the collector voltage and collector current vary according to the broken line drawn in fig. 12. If the scale of v_c is the same as that of $i_c R_2$ (where R_2 is the load resistance in the collector circuit), as is assumed to be the case here, then this line makes an angle of 45° to the axes. The length of this line is proportional to the sweep of i_c and thus a measure for the ratio of that sweep to the sweep of the emitter current: this ratio resembles the mutual conductance of an amplifying tube. With the aid of the apparatus described an oscillogram of this line can easily be produced (*fig. 13*) by recording v_c as a function of i_c with a given alternating current superimposed on the direct current in the emitter circuit.

The length and the position of the line depend, in general, upon the D.C. adjustments of the transistor. Just as it is often desired that the mutual conductance of an amplifying valve varies little

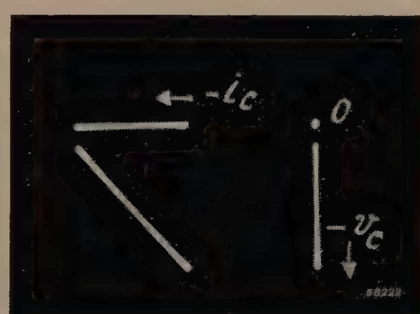


Fig. 13. Collector voltage v_c as a function of the collector current i_c, with an alternating-current source in the emitter circuit.

with the adjustment, so in many cases with a transistor it is desired that the ratio of collector-current sweep to emitter-current sweep varies little with the adjustment. In order to determine the influence

of the collector direct current, this current can be gradually varied by superimposing on the collector current an alternating current with a frequency much lower than that of the emitter current. In

Fig. 14. Oscillogram of the collector voltage as a function of the collector current sweep with a frequency of 50 c/s, while the emitter current sweep with a frequency of 1000 c/s remains within certain limits. The ratio of the two frequencies in this case being exactly a whole number (20), the result is a Lissajous figure.

the oscillogram a series of lines are then traced running at an angle of about 45° (*fig. 14*). To show clearly how this oscillogram is built up, the ratio of the frequencies has been made just equal to a whole number; the oscillogram is then a Lissajous figure. In practice, however, one can work just as well with any ratio of frequencies (provided it is sufficiently large).

The field of characteristics traced in this manner is bounded on the one side by the characteristic $v_c = f(i_c)$ for the highest value of i_e and on the other side by the corresponding characteristic for the lowest value of i_e. As fig. 14 shows, these boundary characteristics run for a part almost parallel and thus the intermediate lines at an angle of 45° are approximately constant in length. Therefore within this area the requirement that the ratio of collector and emitter current sweeps should depend little upon the adjustment is satisfied.

Collector current and collector voltage as functions of the emitter current

As previously mentioned, with the apparatus described it is not necessary that the pairs of terminals from which the voltage differences are to be traced as functions of each other should have one common terminal. To demonstrate this, in *fig. 15* two oscillograms are shown which were taken without any common terminal, as may be seen from the circuit diagram in *fig. 16*: these are recordings of the collector current and of the collector voltage as functions of the emitter current.

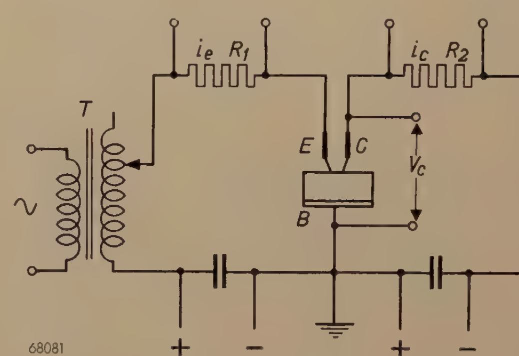


Fig. 16. Circuit for recording the collector current or collector voltage as a function of the emitter current. In one case i_cR_2 and in the other case v_c is recorded as a function of i_eR_1. It should be noted that R_1 has no point in common with the collector circuit.

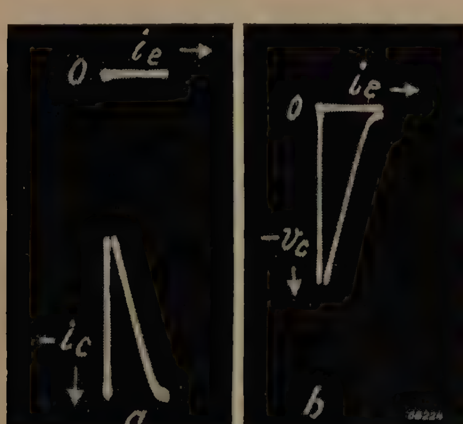


Fig. 15. *a*) Collector current, *b*) collector voltage, both as functions of the emitter current.

The first of these oscillograms shows that within a certain area there is a practically linear relation between i_c and i_e (as also appeared from fig. 14). The second oscillogram shows a remarkable analogy with the characteristic of the anode current as a function of the grid voltage of an amplifying valve.

Transducer gain

In fig. 1 a transistor was represented connected as an amplifier. The voltage at the input is now assumed to be derived from an alternating-voltage source, with an r.m.s. value V_1 and internal resistance R_1. The r.m.s. value of the alternating voltage across the output resistance R_2 will be denoted by V_2.

An important factor is the ratio of the A.C. power developed in R_2 — thus $V_2{}^2/R_2$ — to the maximum output power of the alternating-voltage source V_1, i.e. $(\frac{1}{2}V_1)^2/R_1$. In communication technique this

ratio is called the transducer gain k, which is thus written as:

$$k = \frac{V_2^2/R_2}{(\frac{1}{2}V_1)^2/R_1} \quad \dots \dots \quad (3)$$

This quantity has a maximum for a given value of R_2 and, with the present transistors, may assume a value of about 100.

If R_1 and R_2 are made variable, and their values are known, then, as (3) shows, k can be determined by measuring V_1 and V_2.

V_1' is used as input voltage V_1:

$$V_1 = aV_1' \quad \dots \dots \dots \quad (4)$$

With the aid of the other arm of S_1 the value of R_1 is varied. The tappings on the transformer and on the resistor are so chosen that in the four positions of S_1 the ratio a^2/R_1 remains constant.

The lower arm of a second two-pole switch, S_2, short-circuits an adjustable part of the resistor in the collector circuit, the remaining part being R_2. This short-circuiting takes place across a capacitor

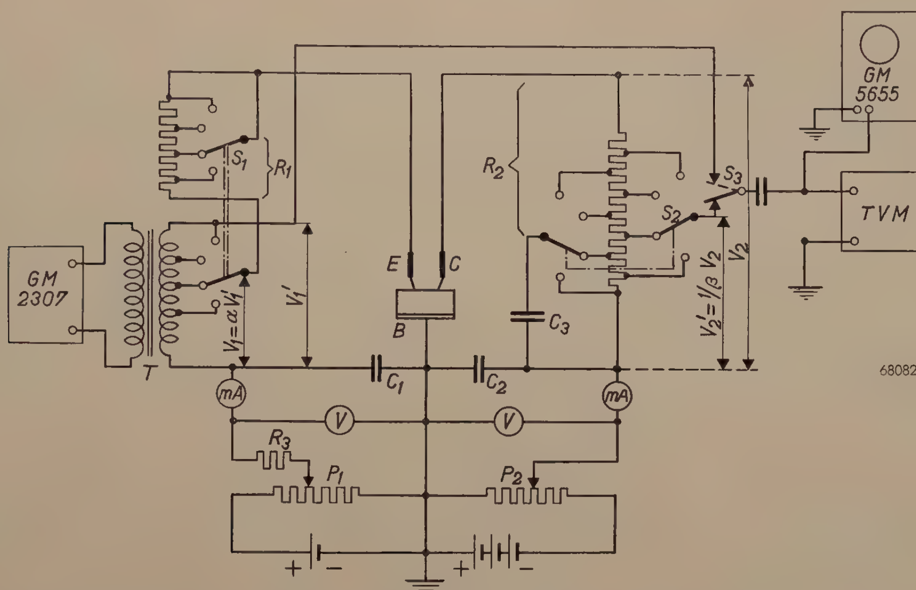


Fig. 17. Circuit for the direct reading of the transducer gain of a transistor (*B-E-C*). *T* input transformer fed from a signal generator GM 2307 adjusted to, say, 1000 c/s. S_1 switch with which a voltage $V_1 = aV_1'$ is tapped from the transformer and at the same time a certain resistance R_1 is included in the emitter circuit. By means of the switch S_2 the value of the resistance R_2 in the collector circuit is adjusted (C_3 is a capacitor to be regarded as a short-circuit for the alternating current) and at the same time the proportion β tapped from the alternating voltage across R_2. The tappings on the transformer and the resistors have been so chosen that $\beta^2/R_2 = 25\ a^2/R_1 =$ constant.

P_1, P_2 potentiometers for controlling the direct-current adjustments. R_3 high resistance maintaining the direct-current adjustment in the emitter circuit when varying R_1. C_1, C_2 bypass capacitors.

Via the switch S_3 the voltage V_1' is compared with V_2' on the valve voltmeter *TVM*, which has db and mW scales. When with V_1' the reading is 20 db, the gain with V_2' is read directly in decibels.

Our measuring set-up is based on this principle, but by an artifice the value of the transducer gain can be read directly in decibels. The circuit is represented in *fig. 17*.

Here, instead of the voltages V_1 and V_2 being measured, two voltages V_1' and V_2', related in a certain manner to V_1 and V_2 respectively, are compared one with the other (with the aid of the switch S_3 and a valve voltmeter). V_1' is the constant alternating voltage of the whole of the secondary winding of a transformer T fed, for instance, from a signal generator adjusted to 1000 c/s. Via the lower arm of a two-pole switch S_1 a proportion a of the voltage

C_3, so that the variation of R_2 does not affect the D.C. adjustment of the collector [8]).

The second arm of S_2 makes contact with a tapping on R_1. The alternating voltage between this tapping and earth is measured with the valve voltmeter and denoted as V_2', i.e.:

$$V_2 = \beta V_2' \quad \dots \dots \dots \quad (5)$$

The two series of tappings on the resistor are so

---

[8]) Any change in the direct-current adjustment of the emitter when varying R_1 is counteracted by a resistor R_3 (fig. 17) of a much higher value than R_1, so that the current through R_3 is practically independent of R_1.

chosen that in the four positions of S_2 the ratio β^2/R_2 is constant and exactly 25 times the constant ratio α^2/E_1.

Substitution of (4) and (5) in (3) yields:

$$k = \frac{(\beta V_2')^2/R_2}{\frac{1}{4}(\alpha V_1')^2 R_1} = 4\,\frac{\beta^2/R_2}{\alpha^2/R_1}\left(\frac{V_2'}{V_1'}\right)^2 = 4 \times 25\left(\frac{V_2'}{V_1'}\right)^2 = 100\left(\frac{V_2'}{V_1'}\right)^2 .$$

If, for instance, we find that V_2' is just equal to V_1' then $k = 100$.

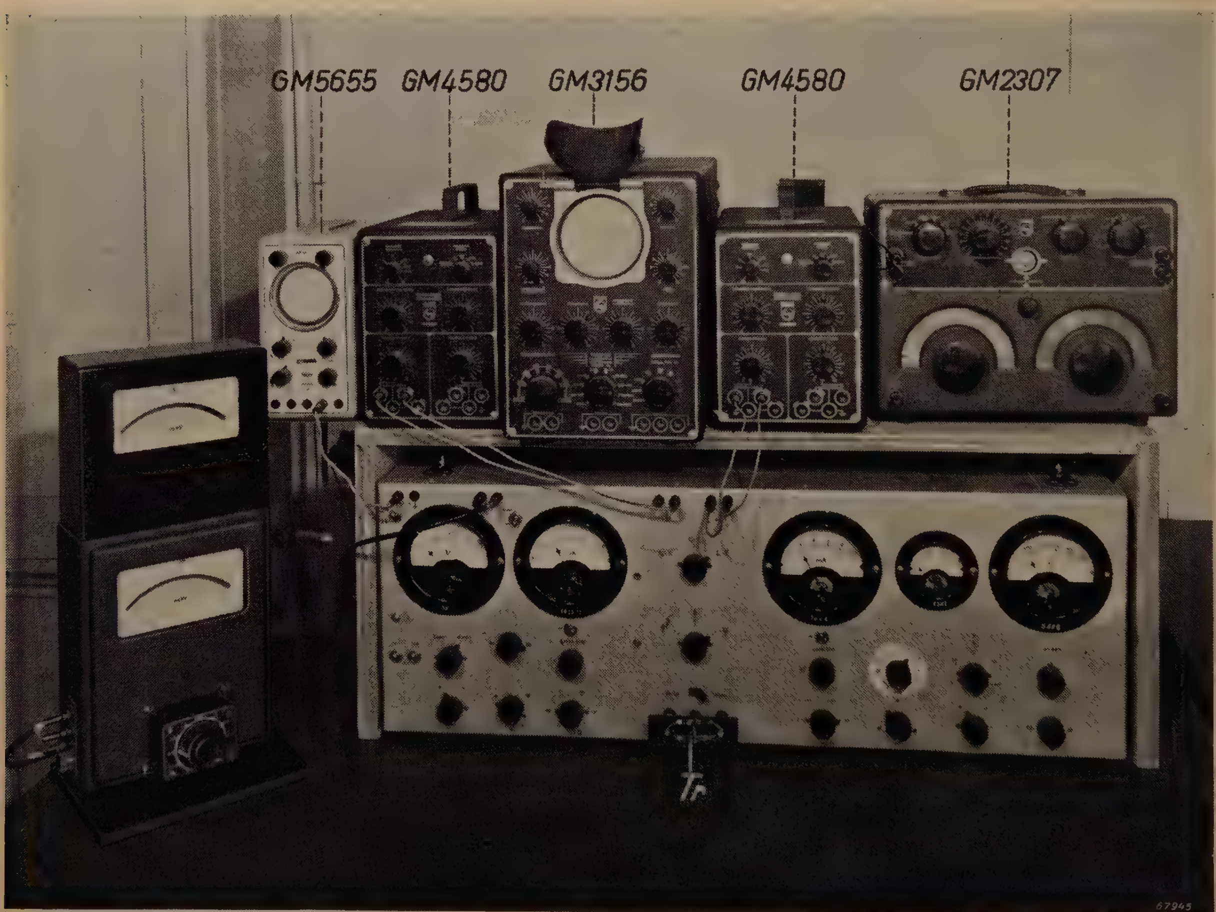


Fig. 18. General picture of the installation. *Tr* transistor under test. To the left and right of the transistor are the controls and meters for the direct-current adjustments and for the resistors in the emitter and collector circuits. Above that, from left to right: checking oscilloscope GM 5655 (cf. fig. 17), electronic switch GM 4580, oscilloscope GM3156, second electronic switch GM 4580, signal generator GM 2307. On the extreme left the valve voltmeter with two moving-coil meters.

When the switch S_3 is put in the position indicated by a broken line the voltage V_1' is applied to the terminals of the valve voltmeter. This meter could be so calibrated as to read 100 at a voltage having a suitable value to serve as V_1' and a value $100\,p^2$ at a voltage of $V_2'=pV_1'$, so that k can be read directly.

We have given preference, however, to readings of k in decibels. As we have seen, $V_2' = V_1'$ corresponds to $k = 100$, i.e. $10 \log 100 = 20$ db, and $V_2 = pV_1'$ thus corresponds to $k = 100\,p^2$, i.e. $10 \log (100\,p^2) = (20 + 20 \log p)$ db. The scale on the voltmeter is such that with a voltage V at the terminals the reading is:

$$c \text{ db} = 20 \log (V/V_0) \text{ db},$$

where V_0 is a voltage chosen as zero level. Thus when measuring V_2' the reading is:

$$a \text{ db} = 20 \log (V_2'/V_0) \text{ db},$$

and when measuring V_1':

$$b \text{ db} = 20 \log (V_1'/V_0) \text{ db}.$$

The difference between these two readings is:

$$a - b = 20 \log (V_2'/V_1') = 20 \log p .$$

Thus the transducer gain to be measured is:

$$k = (20 + a - b) \text{ db}.$$

When V_1' is chosen such that $b = 20$ db, the reading a gives k directly in decibels.

The moving-coil instrument of the valve voltmeter has, in addition to a scale calibrated in decibels, a scale calibrated according to $V_2'^2$, from which can be read in mW the A.C. output developed in R_2. This output is $V_2^2/R_2 = \beta^2 V_2'/R_2 = \text{const.} \times V_2'^2$ so that when the position of the switch S_2 is changed the calibration of the scale remains the same. Another switch (not drawn) provides for the choice of two sensitivities: in one position the full deflection corresponds to 23 db and 1 mW, in the other position it corresponds to 34 db and about 12 mW. Since it is not practicable to have four scales on one instrument, two moving-coil meters are used, each with one decibel and one mW scale. The change-over from one meter to the other is effected together with the selection of the sensitivity.

When choosing V_1 care has to be taken that no distortion arises. As a check upon this, an oscilloscope is connected in parallel with the terminals of the voltmeter (fig. 17). If the oscillogram shows distortion of the output voltage then V_1' has to be reduced until the distortion disappears.

A picture of the whole set-up is given in *fig. 18.*

Summary. The transistor is a circuit element which, like the triode, is capable of amplifying signals. One particular design consists of a germanium crystal (with an excess of conduction electrons), soldered onto a base electrode, and two metal points (the emitter and the collector) placed close together and resting on the crystal.

Here an apparatus is described for tracing various transistor characteristics and measuring the transducer gain.

For the first purpose two electronic switches, type GM 4580, are employed. The emitter voltage, for instance, is applied to one pair of input terminals of one switch, while the voltage produced by the emitter current in a resistor is applied to one pair of input terminals of the other switch. In the emitter circuit there is also an alternating-voltage source. The output terminals of the electronic switches are connected to the vertical and horizontal deflection plates of an oscilloscope tube, which displays the characteristic (emitter voltage as a function of the emitter current) and also — since the other pairs of input terminals of the electronic switches remain open — the coordinate axes and the origin.

It is shown that — thanks to the symmetrical inputs and the method of applying negative feedback in the electronic switches — it is not necessary to have a common terminal for the voltages that are traced one against the other. When the measuring apparatus is connected this does not, therefore, automatically involve earthing of a particular point of the circuit, which, owing to the inevitable stray capacitances, would lead to the formation of a loop on the oscillogram. On the contrary, one is now free to choose the earthing point of the circuit such as to eleminate the effect of the stray capacitances.

In a similar manner it is possible to trace other characteristics, such as the collector voltage as a function of the collector current with a constant or a varying emitter current, or the collector current as a function of the emitter current, etc.

The transducer gain is measured with a different circuit. The gain is indicated by a meter calibrated in decibels.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF THE N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk can be obtained free of charge upon application to the address printed on the back cover.

1962: K. F. Niessen: On the deviations between theoretical and experimental values of the specific heat of superconductors (Physica **16**, 709-718, 1950, No. 9).

The deviations mentioned in the title are explained qualitatively by the combination of two effects, the more important of which is based on the assumption that the thickness of Heisenberg's superconducting layer, covering part of the Fermi surface, does not drop abruptly to zero at the boundary of the layer. The other effect, which was more obvious but appeared to be of less importance, comes into play when the thickness of the superconducting layer decreases with increasing temperature.

1963: J. de Jonge, R. J. H. Alink and R. Dijkstra: Absorption spectrum and photodecomposition of o-hydroxybenzene diazonium sulphate (Rec. Trav. chim. Pays-Bas **69**, 1448-1454, 1950, No. 11).

The long-wave absorption of o-hydroxybenzene-diazonium salts in water is found at 3950 Å, but in 50% sulphuric acid as solvent it appears at 3500 Å. While the photodecomposition of o-hydroxybenzenediazonium sulphate in water leads quantitatively to a derivative of cylopentadiene, decomposition of the same compound in 50% sulphuric acid results for the greater part in catechol. The effect of sulphuric acid, both on the spectrum and the photo-decomposition, may be due to an increase in energy of the quinoid resonance structure of the hydroxybenzenediazonium ion by the acid.

1964: D. Polder: Ferrite materials (Proc. Inst. El. Engrs. **97** II, 246-256, 1950, No. 56).

Survey of electrical and magnetic behaviour of ferrite materials in connection with physico-chemical properties. The manufacture is dealt with and important possibilities for application of different ferrite materials are briefly reviewed. Some of the experimental methods are described for measuring the chemical and magnetic properties in different frequency ranges.

1965*: J. M. Stevels: Expériences et théories sur la facteur de puissance des verres en fonction de leur composition (Verres et Réfractaires **4**, 83-89, 1950, No. 2). (Experiments and theory on power factors of glasses as a function of their composition; in French)

Translated from **R 127**.

1966: J. I. de Jong: Note on the reaction of urea and formaldehyde (Rec. Trav. chim. Pays-Bas, **69**, 1568, 1950, No. 12).

In contradiction to other investigators a purely bimolecular reaction without an initial fall in p_H is found when neutral formol is added to a freshly prepared solution of repeatedly crystallised urea. The decrease in p_H and the high initial reaction rate found by others may be due to small amounts of ammonium salts.

1967: J. D. Fast: Le vieillissement du fer et de l'acier (Revue Métallurgie **47**, 779-786, 1950, No. 10). (Ageing of iron and steel, in French)

For the contents of this paper see Philips Techn. Rev. **13**, 1951 (No. 6).

1968*: Balth. van der Pol and H. Bremmer: Operational calculus based on the two-sided Laplace integral (Cambridge Univ. Press 1950, XIII + 424 pp., 97 figs.).

After a short introduction concerning the person of Oliver Heaviside and his "operational calculus" the Laplace transformation is dealt with at length. Apart from some historical remarks on the one-sided transformation (integration between 0 and ∞) the two-sided Laplace integral ($-\infty$ to $+\infty$) is introduced *ab initio*.

The Fourier integral is dealt with as an introduction to the Laplace transformation. A number of simple examples and elementary rules are given, after which the very essential conceptions of unit function and delta function are introduced. Two chapters are devoted to convergence questions and asymptotic expressions. Then follows among others the treatment of linear differential equations with constant coefficients, *idem* with variable coefficients, step functions and saw tooth functions, differential equations and integral equations. Partial differential equations are dealt with both from the point of the Laplace transformation with one variable and

from that of the "simultaneous" transformation in many variables. The special advantages of the operational calculus come to the fore in deducing unexpected and complex relations between functions and in dealing with transient phenomena in linear networks (filters).

The book, which contains many worked-out examples, is concluded by a list of rules ("grammar") and a list of transformations ("vocabulary").

1969: G. W. Rathenau and J. L. Meijering: Rapid oxydation of metals and alloys in the presence of MoO_3 (Metallurgia **42**, 162-172, 1950, No. 251).

The influence of MoO_3 on the oxidation in air of copper, aluminium bronze, silver, aluminium silver, nickel, nichrome and two heat-resisting steels is measured as a function of temperature. Discontinuities in the curves are correlated with eutectic temperatures measured in mixtures of MoO_3 and oxides of the metals concerned. Metallographic examination of the oxidized specimens reveals some peculiar effects.

Besides by the formation of liquid oxidic phases, the oxidation can be enhanced by MoO_3 in two other ways: incorporation of Mo in the crystal lattice of the oxide skin, and thermodynamic stabilization of Ag_2O by formation of molybdates (cf. also Philips techn. Review **12**, 213-220, 1951, No. 8).

1970: J. Haantjes and F. W. de Vrijer: Flicker in television pictures (Wireless Engineer **28**, 40-42, 1951, No. 2).

See Philips techn. Rev. **13**, 55-60, 1951 (No. 3).

1971: N. W. H. Addink: Quantitive spectrochemical analysis by means of the direct current carbon arc, I. General methods. (Rec. Trav. chim. Pays-Bas **70**, 155-167, 1951, No. 2).

See Philips techn. Rev. **12**, 337-348, 1951 (No. 12).

1972: N. W. H. Addink: Quantitive spectrochemical analysis by means of the direct current carbon arc, II. Biological materials. A possible correlation between the zinc content of liver and blood and the cancer problem (Rec. Trav. chim. Pays-Bas **70**, 168-181, 1951, No. 2).

For the contents of this paper see abstracts No. **1947**.

1973: N. W. H. Addink: The degree of imperfection of crystals (Rec. Trav. chim. Pays-Bas **70**, 202-208, 1951, No. 2).

An attempt is made to arrange crystals of various materials according to their degree of imperfection, derived from the apparent value of N (the Avogadro number). N has been determined for each kind of crystal according to the formula: $N = cM/Vd$, where M is the molecular weight, c a factor related to the number of molecules per unit cell, V the volume of the unit cell and d the density of the crystal. Metals, in accordance with their mosaic structure, show a relatively high value of N and therefore a high degree of imperfection (0.06% of their volume), whilst the lowest value of N has been found for diamond, quartz, calcite and potassium chloride (prepared from a solution), namely $N_{\text{chem}} = (6.0228 \pm 0.0014)\ 10^{23}$, as compared with Cohen's and Dumond's value (6.02378 $\pm$ 0.00011) 10^{23}. Special attention is drawn to the behaviour of perfect crystals of potassium chloride (from a solution) and imperfect crystals of the same substance (from the melt): adsorption phenomena of water vapour enable the thickness of mosaic blocks to be estimated. A value of 6×10^{-4} cm has been found, assuming a monomolecular layer of water molecules on the internal surface.

1974: R. van der Veen: Influence of day-length on the dormancy of some species of the genus Populus (Physiologica Plantarum **4**, 35-40, 1951).

Experiments show that dormancy of poplar trees (Populus) is induced solely by day length. In long days (14 hours and longer) growth and development of new leaves is maintained; in short days (12 hours and shorter) the trees go into dormancy irrespective of the temperature. At 20°C and 29°C all eight species of Populus investigated had stopped their growth in less than 45 short days. After 102 short days all these trees had shed their leaves and had gone into complete dormancy.

1975*: P. J. Bouma: Farbe und Farbwahrnehmung. Einführung in das Studium der Farbreize und Farbempfindungen. (N.V. Philips Gloeilampenfabrieken, Techn. and Scientific Literature Dept. 1951 (XVI + 358 pp., 113 figs.).

German translation of "Physical aspects of colour" by the same author; see Philips techn. Rev. **9**, 158, 1947.

1976*: J. M. Stevels: Expériences et théories sur le facteur de puissance des verres en fonction de leur composition, II (Verres et Réfractaires 5, 4-14, 1951, No. 1). (Some experiments and theories on the power factor of glasses as a function of their composition;in French).

French translation of **R 158**.

R 152: C. J. Bouwkamp: On the diffraction of electromagnetic waves by small circular disks and holes (Philips Res. Rep. 5, 401-422, 1950, No. 6).

This paper deals with the diffraction of a plane-polarized electromagnetic wave by a conducting circular disk for the case of normal incidence. Integro-differential equations are derived for the currents induced in the disk. These equations are approximately solved on the assumption that the radius of the disk is small compared to the wave-length. Six terms of a power-series solution in the basic variable ka are derived, where k is the wave number and a the radius of the disk. The scattered field on the surface of the disk is calculated, to the same degree of accuracy, and also the field in the wave zone. An expression is obtained for the scattering coefficient of the disk. Finally, Babinet's principle is applied to obtain the corresponding solution for the diffraction of a plane wave by a circular hole in an infinite, plane, conducting screen (see **R 148**).

R 153: G. Diemer: Passive feedback admittance of disc-seal triodes (Philips Res. Rep. 5, 423-434, 1950, No. 6).

At microwave-frequencies the inductance of the grid wires in disc-seal grounded-grid triodes plays an important part in the passive feedback from anode to cathode. It is shown that with a well-designed valve geometry this can be used to neutralize more or less the capacitive feedback through the anode-to-cathode capacitance.

R 154: J. M. L. Janssen: The method of discontinuities in Fourier analysis (Philips Res. Rep. 5, 435-460, 1950, No. 6).

A survey is given of the results arrived at in various publications dealing with the method of discontinuities in Fourier analysis. It is shown that a more general conception is practicable when starting from the Fourier integrals instead of the Fourier series. Then in a simple and natural way also the frequency spectrum of continuous functions may be found by this method. It is investigated in how far rapid changes may be construed as being real discontinuities. Correction factors are derived which take into account the shape of the rapid changes. Several examples are given, whilst, *inter alia*, the method is applied to the problem of the summation of a series, an Euler summation formula then being found.

R 155: W. Ch. van Geel and B. C. Bouma: Variation en fonction de la fréquence de la caractéristique dynamique $i(V)$ du système Al-Al_2O_3-électrolyte (Philips Res. Rep. 5, 461-475, 1950, No. 6). (Variation of the dynamic characteristic $i(V)$ of the system Al-Al_2O_3-electrolyte as a function of the frequency; in French.)

The current-voltage characteristic of the system Al-Al_2O_3-electrolyte is not a single curve but shows a loop. With increasing frequency the height of the loop (maximum current) decreases. At about 5000 c/s the loop has almost disappeared and the system has lost its rectifying properties. For a newly formed layer, the current in the direction of easy flow increases strongly during the first moments after the application of an alternating voltage. Some seconds are necessary for the current to reach its maximum. The constitution of the layer is not a stable one. For rectification a particular constitution in the layer has to be created and after that the situation is still not stable but changes with the direction of the applied voltage. An effort is made to explain the phenomena observed. It is assumed that the oxide layer contains an excess of aluminium in that part of the layer which is bounded by the aluminium. This part is semi-conducting (excess semi-conductor). The other part of the layer is the barrier layer, the thickness of which changes with the direction of the applied voltage and gives rise to the loop. The variation of the thickness is caused by electrolysis in the Al_2O_3 layer.

Another possibility is to assume that the part of the oxide layer near the electrolyte contains a surplus of oxygen and forms a defective semi-conductor.